

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2003, 14:45:15 ; Search time 15.7768 Seconds  
(without alignments)  
1713.885 Million cell updates/sec

Title: US-09-854-356-6

Perfect score: 5078  
Sequence: 1 MEALALCRWGLLALLPGAA.....TFKGTPTAENPEYLGDVYV 919

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCCTS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/Dackfilest1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4900	96.5	1255	2	US-08-625-101-2
2	4900	96.5	1255	2	US-08-336-786-2
3	4891	96.3	1255	1	US-08-467-083-68
4	4891	96.3	1255	1	US-08-414-417B-68
5	4891	96.3	1255	2	US-08-486-348A-68
6	4891	96.3	1255	2	US-08-468-545B-68
7	4891	96.3	1255	3	US-08-466-680B-68
8	4818	94.9	1255	2	US-08-484-438-8
9	3632	71.5	782	2	US-09-146-283-4
10	3632	71.5	782	3	US-08-579-823A-4
11	3632	71.5	782	4	US-09-344-195-4
12	3473	68.4	624	4	US-08-422-108-1
13	3473	68.4	624	4	US-08-422-734-1
14	1878	37.0	419	4	US-09-630-155-2
15	1645	32.4	1210	2	US-08-464-438-7
16	1645	32.4	1210	2	US-08-464-438-7
17	1533	30.2	644	1	US-08-475-035-4
18	1528.5	30.1	1308	2	US-08-336-708A-9
19	1474	29.0	580	1	US-08-414-417B-69
20	1474	29.0	580	2	US-08-486-348A-69
21	1474	29.0	580	2	US-08-486-348A-69
22	1474	29.0	580	3	US-08-466-680B-69
23	1466.5	28.9	911	2	US-08-484-438-10
24	1451	28.3	1058	2	US-08-484-438-4
25	1435	28.3	1342	1	US-07-978-895-4
26	1435	28.3	1342	1	US-08-484-438-4
27	1435	28.3	1342	2	US-08-473-119-4

28	1435	28.3	1342	2	US-08-475-352-4	Sequence 4, Appl1
29	1427.5	28.1	1343	6	5183884-4	Patent No. 5183884
30	1028.5	20.3	478	4	US-09-570-454-2	Sequence 2, Appl1
31	493	9.7	97	1	US-08-421-356-3	Sequence 3, Appl1
32	493	9.7	97	4	US-09-046-783-3	Sequence 3, Appl1
33	368	7.2	1382	2	US-08-737-715-2	Sequence 2, Appl1
34	368	7.2	1382	4	US-09-457-040B-7	Sequence 2, Appl1
35	313.5	6.2	1367	2	US-08-625-819-2	Sequence 2, Appl1
36	311.5	6.1	516	3	US-08-746-559A-4	Sequence 2, Appl1
37	311.5	6.1	1367	2	US-08-249-687C-2	Sequence 2, Appl1
38	311.5	6.1	1367	3	US-08-746-559A-2	Sequence 2, Appl1
39	305.5	6.0	705	2	US-08-456-647B-4	Sequence 2, Appl1
40	305.5	6.0	705	2	US-08-237-401A-4	Sequence 2, Appl1
41	303.5	6.0	1367	4	US-08-864-641B-18	Sequence 4, Appl1
42	295.5	5.8	486	3	US-08-746-559A-5	Sequence 18, Appl1
43	291	5.7	370	4	US-08-857-076-104	Sequence 5, Appl1
44	269.5	5.3	541	2	US-08-484-438-6	Sequence 104, App
45	267.5	5.3	1724	4	US-08-857-076-12	Sequence 6, Appl1
						Sequence 12, Appl1

## ALIGNMENTS

```
RESULT 1
US-08-625-101-2
; Sequence 2, Application US/08625101
; Patent No. 5869445
;
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disis, Mary L.
; TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE
; TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION
; TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/625,101
; FILING DATE: 01-Apr-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-625-101-2
```

```
Query Match 96.5%; Score 4900; DB 2; Length 1255;
Best Local Similarity 73.2%; Pred. No. 0;
Matches 919; Conservative 0; Mismatches 0; Indels 336; Gaps 1;
CY 1 MEALALCRWGLLALLPGAASTQVCTGDMKRLRPASPEETHLDMRLHYGCGQVVGNTL 60
DB 1 MEALALCRWGLLALLPGAASTQVCTGDMKRLRPASPEETHLDMRLHYGCGQVVGNTL 60
```

```

QY 61 ELTYLPTNASLSTLQDIOEVQGVLLAHNOVROVPLQRLRIYRGTOLEFEDNALAVLNDG 120
DB 61 ELTYLPTNASLSTLQDIOEVQGVLLAHNOVROVPLQRLRIYRGTOLEFEDNALAVLNDG 120
QY 121 DPLNNTPTVGTASPGGLRELOLRSLEILKGVLLIORNPOLCYODTILMKDIFHKNNOLA 180
DB 121 DPLNNTPTVGTASPGGLRELOLRSLEILKGVLLIORNPOLCYODTILMKDIFHKNNOLA 180
QY 181 LTLIDNRSRACHPCSPMCKSGRCWGESSEDCOSTLRTVCAGGACARCKGFLPTDCCHEQC 240
DB 181 LTLIDNRSRACHPCSPMCKSGRCWGESSEDCOSTLRTVCAGGACARCKGFLPTDCCHEQC 240
QY 241 AACCTGPKHSDCLACILAHNHSIGCEILHCPALVYNTDTEFSMNPGRRTYFGASCTYACP 300
DB 241 AACCTGPKHSDCLACILAHNHSIGCEILHCPALVYNTDTEFSMNPGRRTYFGASCTYACP 300
QY 301 YNLSTDVGSCTLVCPALHNOEVAEDGTQRCCKSKPCARVCYGLGMEHLREYRAVTSAN 360
DB 301 YNLSTDVGSCTLVCPALHNOEVAEDGTQRCCKSKPCARVCYGLGMEHLREYRAVTSAN 360
QY 361 IOEFAGCKITFGSLAFLPESFPDDPASNTAPLOPBOLOVETLEETLGYLYISAMPDLSL 420
DB 361 IOEFAGCKITFGSLAFLPESFPDDPASNTAPLOPBOLOVETLEETLGYLYISAMPDLSL 420
QY 421 DLSVEONLOVIRGRILHNGAYSLTLOGLSIMLGRSLRELGSGLALIHNTHLCEVHTV 480
DB 421 DLSVEONLOVIRGRILHNGAYSLTLOGLSIMLGRSLRELGSGLALIHNTHLCEVHTV 480
QY 481 PMDOLFNNPQALHTANRPDECEVGEGLACIOLCARGHCGMPPTOCVNCOSFLRGQEC 540
DB 481 PMDOLFNNPQALHTANRPDECEVGEGLACIOLCARGHCGMPPTOCVNCOSFLRGQEC 540
QY 541 VEECRVILQGLPREVYNARHCLPCHEQOPQNGSTTCGPRADQVCAAHKDPFCVARGC 600
DB 541 VEECRVILQGLPREVYNARHCLPCHEQOPQNGSTTCGPRADQVCAAHKDPFCVARGC 600
QY 601 PSQVPRDLSYMPKFPDEBEGACOPCPLNCTHSCVDLDDKCPAEORASPLTS----- 653
DB 601 PSQVPRDLSYMPKFPDEBEGACOPCPLNCTHSCVDLDDKCPAEORASPLTS----- 653
QY 654 ----- 653
DB 654 ----- 653
QY 661 ILLVVLGVNFGILIKRROQKIRKRYTMRLLQETELVEPLTPSCAMPNOAMRIKETEL 720
DB 661 ILLVVLGVNFGILIKRROQKIRKRYTMRLLQETELVEPLTPSCAMPNOAMRIKETEL 720
QY 654 ----- 653
DB 654 ----- 653
QY 721 RKVYLGSGAGTGYKGIWIPDGENVKIPVAKVIRENTSPKANKELIDEAYNAGVGP 780
DB 721 RKVYLGSGAGTGYKGIWIPDGENVKIPVAKVIRENTSPKANKELIDEAYNAGVGP 780
QY 654 ----- 653
DB 654 ----- 653
QY 781 YVSRLLGICLSTYQVLTQMLPYGCLLDHVHENGRLGSQLLMKMQIAKMSLYLDDR 840
DB 781 YVSRLLGICLSTYQVLTQMLPYGCLLDHVHENGRLGSQLLMKMQIAKMSLYLDDR 840
QY 654 ----- 653
DB 654 ----- 653
QY 841 LVHRDLAARNVLYKSPNHVKITDFGLARLIDIDEYHADGKVPKIMMALESILRRFT 900
DB 841 LVHRDLAARNVLYKSPNHVKITDFGLARLIDIDEYHADGKVPKIMMALESILRRFT 900
QY 654 ----- 653
DB 654 ----- 653
QY 901 HQSDVWSYGVYVWELMTFGANPYDGIIPAREIPDLLEKGERLPORPCTIIVYMIWKCM 960
DB 901 HQSDVWSYGVYVWELMTFGANPYDGIIPAREIPDLLEKGERLPORPCTIIVYMIWKCM 960
QY 654 ----- 653
DB 654 ----- 653
QY 961 IDSECRPFRELFVSEFMSARDPQRFVYIQNEDLGPASPLDSTFRTSLLEDDMDLDA 1020
DB 961 IDSECRPFRELFVSEFMSARDPQRFVYIQNEDLGPASPLDSTFRTSLLEDDMDLDA 1020
QY 685 BEYLVPQGFPCPDPAAGAGVHHRRSSSTRSGGGLTJGLEPSEEDAPRSLAPSEG 744
DB 685 BEYLVPQGFPCPDPAAGAGVHHRRSSSTRSGGGLTJGLEPSEEDAPRSLAPSEG 744
QY 1021 BEYLVPQGFPCPDPAAGAGVHHRRSSSTRSGGGLTJGLEPSEEDAPRSLAPSEG 1080
DB 1021 BEYLVPQGFPCPDPAAGAGVHHRRSSSTRSGGGLTJGLEPSEEDAPRSLAPSEG 1080
QY 745 AGSDVFDGDLGMAKGIQSLPTHPSPLQORSSEDPYVPLPSMTGCVYAPLTCSPQRYV 804
DB 745 AGSDVFDGDLGMAKGIQSLPTHPSPLQORSSEDPYVPLPSMTGCVYAPLTCSPQRYV 804
QY 1081 AGSDVFDGDLGMAKGIQSLPTHPSPLQORSSEDPYVPLPSMTGCVYAPLTCSPQRYV 1140
DB 1081 AGSDVFDGDLGMAKGIQSLPTHPSPLQORSSEDPYVPLPSMTGCVYAPLTCSPQRYV 1140

```

```

QY 805 NOPVAPPPSPREGPLPAPAPAGATLERKTLSPGKNGVKDVAFGAVENPEYLTPQ 864
DB 1141 NOPVAPPPSPREGPLPAPAPAGATLERKTLSPGKNGVKDVAFGAVENPEYLTPQ 1200
QY 865 GGAAPQPPPPAFSPAFDNLTYWDDPPRGRAPSTFKGTPTAENPEYLGLDVPY 919
DB 1201 GGAAPQPPPPAFSPAFDNLTYWDDPPRGRAPSTFKGTPTAENPEYLGLDVPY 1255

RESULT 2
US-08-356-786-2
; Sequence 2, Application US/08356786
; Patent No. 5877305
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppertmann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,786
FILING DATE: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/831,967
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: CRP-053
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ. ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-356-786-2

Query Match 96.58; Score 4900; DB 2; Length 1255;
Best Local Similarity 73.2%; Pred. No. 0;
Matches 919; Conservative 0; Mismatches 0; Indels 336; Gaps 1;

```

Db 181 LTLIDTNRSRACHPCSPMKGSKRCWGESSESDCSLRTTVCAGGCARCKGPTDCCHEOC 240  
QY 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDTFESMPNBEGRYFGASCYACR 300  
Db 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDTFESMPNBEGRYFGASCYACR 300  
QY 301 YNTLSTDVSGCTLVCPHLHNOEYTAEDGTORCEKSKPCARVCYGLGMEHLREVRATYSAN 360  
Db 301 YNTLSTDVSGCTLVCPHLHNOEYTAEDGTORCEKSKPCARVCYGLGMEHLREVRATYSAN 360  
QY 361 IOEFACCKRIFGSLAFPESEFDGDPASNTAPLOPEOLQVETLEITGYLIYSAMPDLSL 420  
Db 361 IOEFACCKRIFGSLAFPESEFDGDPASNTAPLOPEOLQVETLEITGYLIYSAMPDLSL 420  
QY 421 DLSVPQNLQYIRGRILLHNGAYSTTLOGGISMGLSLRGLSGGLAIHNTHLCEVHTV 480  
Db 421 DLSVPQNLQYIRGRILLHNGAYSTTLOGGISMGLSLRGLSGGLAIHNTHLCEVHTV 480  
QY 481 PMDQFRNPQHALLHTANRPEDCEVGEGLACHOLCARGHGMGPPTQVCVNSQFLRGQEC 540  
Db 481 PMDQFRNPQHALLHTANRPEDCEVGEGLACHOLCARGHGMGPPTQVCVNSQFLRGQEC 540  
QY 541 VEECRVYLOGLPREYVNAHRCLPCHPECOPONGSVTCFGEPAQCVACAHYKDPFCVARC 600  
Db 541 VEECRVYLOGLPREYVNAHRCLPCHPECOPONGSVTCFGEPAQCVACAHYKDPFCVARC 600  
QY 601 PSQVKNDSLTPMKFPDEGACQPCPINCTHSCVDLDKGCACBQASPLTS----- 653  
Db 601 PSQVKNDSLTPMKFPDEGACQPCPINCTHSCVDLDKGCACBQASPLTSIIISAVG 660  
QY 654 ----- 653  
Db 661 ILLVVLGVVFGILLIKRROOKIKKTYMRLLDTELEVEPLTBSGAMPNOAKRILKETEL 720  
QY 654 ----- 653  
Db 721 RKVKVLGSGAFGTVYKGIWLPDGENVKIPAIKVLRENTSPKANKELDEAYVAGVSP 780  
QY 654 ----- 653  
Db 781 YVSRLLIGLTSTVOLVTLMPYCLLDHYRENRGLSGODLLMCMQIAKMSYLEDYR 840  
QY 654 ----- 653  
Db 841 LVHRDLAARNVLKSPNHVKITDFGLARLLDIDETRYHADGCKVPIKMALESILRRFT 900  
QY 654 ----- 653  
Db 901 HOSDVMYGVTWELMTFGAKPYDIPAREIPDLLEKGERLPQPICTIDVYIMVYKMW 960  
QY 654 ----- QNEDLGASPLDSTFYSLLEDDMGDLVA 684  
Db 961 IDSECRPRRELVSERSMARDQRFVYIIONEDLGASPLDSTFYSLLEDDMGDLVA 1020  
QY 685 EETIVYQGGFFCPDPAPGAGWVHHRRSSSTRSGGDLTLGLEPSEEAERAPSLAPSEG 744  
Db 1021 EETIVYQGGFFCPDPAPGAGWVHHRRSSSTRSGGDLTLGLEPSEEAERAPSLAPSEG 1080  
QY 745 AGSDVFDGLGMAKGLDLPHTDPSPLORYSEDPVLPSETDGVVAPLTCSPQPEVY 804  
Db 1081 AGSDVFDGLGMAKGLDLPHTDPSPLORYSEDPVLPSETDGVVAPLTCSPQPEVY 1140  
QY 805 NOPDVAPQPSPREGRLPARAGATLERPKTLSPGKNVYVDFVAFGAVENPEVLTQ 864  
Db 1141 NOPDVAPQPSPREGRLPARAGATLERPKTLSPGKNVYVDFVAFGAVENPEVLTQ 1200  
QY 865 GGAAPQPHPPAPSPAFDNLVYWDODPRPGAPPTSTFKGTPTAENEYELGLDVPV 919  
Db 1201 GGAAPQPHPPAPSPAFDNLVYWDODPRPGAPPTSTFKGTPTAENEYELGLDVPV 1255

RESULT 3  
US-08-467-083-68

; Sequence 68, Application US/08467083  
; Patent No. 5726023  
; GENERAL INFORMATION:  
; APPLICANT: Cheever, Martin A.  
; APPLICANT: Disis, Mary L.  
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/NEU PROTEIN  
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
; NUMBER OF SEQUENCES: 68  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed and Berry  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: US  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/467,083  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/414,417  
; FILING DATE: 06-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharkey, Richard G.  
; REGISTRATION NUMBER: 32,629  
; REFERENCE/DOCKET NUMBER: 920010.448C2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; TELEX: 3723836 SEEDANDBERY  
; INFORMATION FOR SEQ ID NO: 68:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1255 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-467-083-68  
  
Query Match 96.3%; Score 4891; DB 1; Length 1255;  
Best Local Similarity 73.1%; Pred. No. 0;  
Matches 918; Conservative 0; Mismatches 1; Indels 336; Gaps 1;  
  
QY 1 METALCRNGILLALALPPGAASTOVCTGDMKRLRLPASPETHLDMRLHYOGCOVYQGNL 60  
Db 1 METALCRNGILLALALPPGAASTOVCTGDMKRLRLPASPETHLDMRLHYOGCOVYQGNL 60  
QY 61 ELTYLPTNALSFLQDIOEVQGYVLIHNOVROVPLQRLRIYVGTOLLEFDNALAVLDNG 120  
Db 61 ELTYLPTNALSFLQDIOEVQGYVLIHNOVROVPLQRLRIYVGTOLLEFDNALAVLDNG 120  
QY 121 DPLNNTPTVGTASPGGLRELQRLSLREILKGVLTIORNPOLCYODTILKKDIFHKNOA 180  
Db 121 DPLNNTPTVGTASPGGLRELQRLSLREILKGVLTIORNPOLCYODTILKKDIFHKNOA 180  
QY 181 LTLIDTNRSRACHPCSPMKGSKRCWGESSESDCSLRTTVCAGGCARCKGPTDCCHEOC 240  
Db 181 LTLIDTNRSRACHPCSPMKGSKRCWGESSESDCSLRTTVCAGGCARCKGPTDCCHEOC 240  
QY 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDTFESMPNBEGRYFGASCYACR 300  
Db 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDTFESMPNBEGRYFGASCYACR 300  
QY 301 YNTLSTDVSGCTLVCPHLHNOEYTAEDGTORCEKSKPCARVCYGLGMEHLREVRATYSAN 360  
Db 301 YNTLSTDVSGCTLVCPHLHNOEYTAEDGTORCEKSKPCARVCYGLGMEHLREVRATYSAN 360  
QY 361 IOEFACCKRIFGSLAFPESEFDGDPASNTAPLOPEOLQVETLEITGYLIYSAMPDLSL 420  
Db 361 IOEFACCKRIFGSLAFPESEFDGDPASNTAPLOPEOLQVETLEITGYLIYSAMPDLSL 420

```

Db 361 IOEFAGCKKIFGSLAFLEPSEFDGDPASNTAPLQPEQLOVFEETLEITGYLYISAMPDLSL 420
Qy 421 DLSVFQNTLOVINGRILHNGAVSLTLQGLISWLGRLSLRELSGSLALHHNTHLCFVHTV 480
Db 421 DLSVFQNTLOVINGRILHNGAVSLTLQGLISWLGRLSLRELSGSLALHHNTHLCFVHTV 480
Qy 481 PMDOLFRRNPQHALLHTANRPEDECVGEGLAGHQLCARHCWGPGPTQCVNCSQFLRGQEC 540
Db 481 PMDOLFRRNPQHALLHTANRPEDECVGEGLAGHQLCARHCWGPGPTQCVNCSQFLRGQEC 540
Qy 541 VEECVLQGLPREYVNAHRCPLCHPECOPOGNSVTCFGEADQCVACAHYKRPCCVARC 600
Db 541 VEECVLQGLPREYVNAHRCPLCHPECOPOGNSVTCFGEADQCVACAHYKRPCCVARC 600
Qy 601 PSGVKPDLSYMPIMKFPDEGACQPCPINCHTSCVDLDDKGCAPARQASPLTS----- 653
Db 601 PSGVKPDLSYMPIMKFPDEGACQPCPINCHTSCVDLDDKGCAPARQASPLTSIISAVYG 660
Qy 654 ----- 653
Db 661 ILLVVVLGVFGILIKRROOKIRKYTMRLLOETELVEPLPPSGAMPNQAQMRILKETEL 720
Qy 654 ----- 653
Db 721 RKVKVLGSGAGTYKKGIMIPDGENVKIPVAILKREMTSPKANKELIDEAYVMAGVGS 780
Qy 654 ----- 653
Db 781 YVSRLLGICLSTVQLVQLMPYGLLDHVRENRRGLSQDILLNMQIANKMSYLEDR 840
Qy 654 ----- 653
Db 841 LVHRDLAARNVLYKSPNHVKITDFGLARLLDIDETEHADGKVPKWMALESIILRRFT 900
Qy 654 ----- 653
Db 901 HOSDWMSYGVYVWELMTGAKRPGYDIPAREIPDLLEKGBRLPQPICTIDVYMIWKCMM 960
Qy 654 -----ONEDLGPAFLDSTFYRSLLEDDDMKDLYDA 684
Db 961 IDSECRPRFRELVEFSNMARDPQRFVYIQNEDLGPAFLDSTFYRSLLEDDDMKDLYDA 1020
Qy 685 EBYLVPOGFCPCPPAPGAGVHHRRSSSTRSGGDLTLCLEPSEEBAPSPPLAPSEG 744
Db 1021 EBYLVPOGFCPCPPAPGAGVHHRRSSSTRSGGDLTLCLEPSEEBAPSPPLAPSEG 1080
Qy 745 AGSDVFDGDLGAKGKIOSLPTHDSPLOKSEDPVPLPSETGYYAPLTCSPOPEYV 804
Db 1081 AGSDVFDGDLGAKGKIOSLPTHDSPLOKSEDPVPLPSETGYYAPLTCSPOPEYV 1140
Qy 805 NOPDVRRPQPSRREGPLPAARPAAGATLERPKTISPGKNGVYDVAFGAVENPEYLFQ 864
Db 1141 NOPDVRRPQPSRREGPLPAARPAAGATLERPKTISPGKNGVYDVAFGAVENPEYLFQ 1200
Qy 865 GGAPQHPHPPAFPAFNLVWDODPPRCAPSPSTFKGTPTAENPEYIGLDV 919
Db 1201 GGAPQHPHPPAFPAFNLVWDODPPRCAPSPSTFKGTPTAENPEYIGLDV 1255

```

## RESULT 4

US-08-414-417B-68  
Sequence 68, Application US/08414417B  
Patent No. 5801005

GENERAL INFORMATION:  
APPLICANT: Cheever, Martin A.

APPLICANT: Disis, Mary L.  
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN

TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed and Berry LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue

```

: CITY: Seattle
: STATE: Washington
: COUNTRY: US
: ZIP: 98104-7092
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/414,417B
: FILING DATE: 31-MAR-1995
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Sharkey, Richard G.
: REGISTRATION NUMBER: 32,629
: REFERENCE/DOCKET NUMBER: 920010.448C2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 622-4900
: TELEFAX: (206) 682-6031
: INFORMATION FOR SEQ ID NO: 68:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1255 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
:
: US-08-414-417B-68
:
: Query Match 96.3%; Score 4891; DB 1; Length 1255;
: Best Local Similarity 73.1%; Pred. No. 0; Mismatches 1; Indels 336; Gaps 1;
: Matches 918; Conservative 0;
:
: Qy 1 MELAALCRWGLLALLPPGAASVQCTGDMKLRLPASPTHLDMLRHLQGCQVYQNTL 60
: Db 1 MELAALCRWGLLALLPPGAASVQCTGDMKLRLPASPTHLDMLRHLQGCQVYQNTL 60
:
: Qy 61 ETTYLPTNASLSFLDDIQEVGVVLIANQVAVPLQRLRIYRGQLTEEDNTALAVLNG 120
: Db 61 ETTYLPTNASLSFLDDIQEVGVVLIANQVAVPLQRLRIYRGQLTEEDNTALAVLNG 120
:
: Qy 121 DPLNTTPYTGASPGGRELQRLSLTEILKGVLIQNRNPQLCYDPTILMKDIFHKNNOLA 180
: Db 121 DPLNTTPYTGASPGGRELQRLSLTEILKGVLIQNRNPQLCYDPTILMKDIFHKNNOLA 180
:
: Qy 181 LTLIDTNRSRACHPCSPMKSGRSGEESSEDCSLTRVCAGAGCARCKGRLPTDCCHQC 240
: Db 181 LTLIDTNRSRACHPCSPMKSGRSGEESSEDCSLTRVCAGAGCARCKGRLPTDCCHQC 240
:
: Qy 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMPNBEGRITFGASCYTACP 300
: Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMPNBEGRITFGASCYTACP 300
:
: Qy 301 VYVLSIDVSGCTLVCPRLHNOEYTAEDGTORCEKSPCARVYCGMEHLREVRVTSAN 360
: Db 301 VYVLSIDVSGCTLVCPRLHNOEYTAEDGTORCEKSPCARVYCGMEHLREVRVTSAN 360
:
: Qy 361 IOEFAGCKKIFGSLAFLEPSEFDGDPASNTAPLQPEQLOVFEETLEITGYLYISAMPDLSL 420
: Db 361 IOEFAGCKKIFGSLAFLEPSEFDGDPASNTAPLQPEQLOVFEETLEITGYLYISAMPDLSL 420
:
: Qy 421 DLSVFQNTLOVINGRILHNGAVSLTLQGLISWLGRLSLRELSGSLALHHNTHLCFVHTV 480
: Db 421 DLSVFQNTLOVINGRILHNGAVSLTLQGLISWLGRLSLRELSGSLALHHNTHLCFVHTV 480
:
: Qy 481 PMDOLFRRNPQHALLHTANRPEDECVGEGLAGHQLCARHCWGPGPTQCVNCSQFLRGQEC 540
: Db 481 PMDOLFRRNPQHALLHTANRPEDECVGEGLAGHQLCARHCWGPGPTQCVNCSQFLRGQEC 540
:
: Qy 541 VEECVLQGLPREYVNAHRCPLCHPECOPOGNSVTCFGEADQCVACAHYKRPCCVARC 600
: Db 541 VEECVLQGLPREYVNAHRCPLCHPECOPOGNSVTCFGEADQCVACAHYKRPCCVARC 600
:
: Qy 601 PSGVKPDLSYMPIMKFPDEGACQPCPINCHTSCVDLDDKGCAPARQASPLTS----- 653
: Db 601 PSGVKPDLSYMPIMKFPDEGACQPCPINCHTSCVDLDDKGCAPARQASPLTS----- 653

```



Db 601 PSQVPRDLSYMPIMKFPDEEGACQPCPINCTHSCVDLDKGCFAEQRASPLTSLISAVYG 660  
 QY 654 ----- 653  
 Db 661 ILLVVLGVVFGILLIRROOKIRKYMRLLOETELVEPLTPSGAMPNOQMRLKETEL 720  
 QY 654 ----- 653  
 Db 721 RKVKVLGSGAFGVYKGIWIPDGENVKIPVAIKVIRENTSPKANKELDEAYMAGVGSF 780  
 QY 654 ----- 653  
 Db 781 YVSRLLGICLTSTVOLVTOQMPYGCLLDHVRENRRGLSGODLLNMCQIAKMSYLEDR 840  
 QY 654 ----- 653  
 Db 841 LVHRDLAARNVLKSPNHVKITDFGLARLLIDETEHADGKVPITKMALESILRRPT 900  
 QY 654 ----- 653  
 Db 901 HOSDWSYGVYTWELMTFGAKPRPDGIPAREIPDLLEKGERLPOPICTIDVIMYKCM 960  
 QY 654 ----- 684  
 Db 961 IDSECRPRRELVSERSRMAHQRFVVIQNEDLGPASPLDSTFYRSLLEDDMDGLVDA 1020  
 QY 685 EEVLVQOQFPCPDPAFGAGVHNRHSSSTRSGGDLTLGLEPSEEAAPRSLAPSG 744  
 Db 1021 EEVLVQOQFPCPDPAFGAGVHNRHSSSTRSGGDLTLGLEPSEEAAPRSLAPSG 1080  
 QY 745 AGSDVFDGLGMAKAGLSLPHDPSPIQRYSEDPVLPSETPDGVVPLTCSPOPEYV 804  
 Db 1081 AGSDVFDGLGMAKAGLSLPHDPSPIQRYSEDPVLPSETPDGVVPLTCSPOPEYV 1140  
 QY 805 NOPDVRPQPPSPREGPLPAARPAATLERPKTLPKNGVYKDVFAFGAVENPEYLTPQ 864  
 Db 1141 NOPDVRPQPPSPREGPLPAARPAATLERPKTLPKNGVYKDVFAFGAVENPEYLTPQ 1200  
 QY 865 GGAAPQHPHPAPSPAFNDLYWDDPBRGAPSPSTFKTPTAENEYGLDVPV 919  
 Db 1201 GGAAPQHPHPAPSPAFNDLYWDDPBRGAPSPSTFKTPTAENEYGLDVPV 1255  
 RE- 5  
 US-08-486-348A-68  
 Sequence 68, Application US/08486348A  
 Patent No. 3846358  
 GENERAL INFORMATION:  
 APPLICANT: Cheever, Martin A.  
 APPLICANT: Disis, Mary L.  
 TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN  
 TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
 NUMBER OF SEQUENCES: 69  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Seed and Berry LLP  
 STREET: 6300 Columbia Center, 701 Fifth Avenue  
 CITY: Seattle  
 STATE: Washington  
 COUNTRY: US  
 ZIP: 98104-7092  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/486,348A  
 FILING DATE: 07-JUN-1995  
 CLASSIFICATION: 424  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Sharkey, Richard G.  
 REGISTRATION NUMBER: 32,629

REFERENCE/DOCKET NUMBER: 920010.4486  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 622-4900  
 TELEFAX: (206) 682-6031  
 INFORMATION FOR SEQ ID NO: 68:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1255 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 US-08-486-348A-68  
 Query Match 96.3%; Score 4891; DB 2; Length 1255;  
 Best Local Similarity 73.1%; Pred No. 0;  
 Matches 918; Conservative 0; Mismatches 1; Indels 336; Gaps 1;  
 Db 1 MELAALCRWGLLALLPFGAASSTOYCTGDMKRLRPAETHLMDLRHLHYOGQOVVGNL 60  
 1 MELAALCRWGLLALLPFGAASSTOYCTGDMKRLRPAETHLMDLRHLHYOGQOVVGNL 60  
 QY 61 ELTYLPTNASLFLDIOIEYQGVYVLAHQVQVPLQRLRYNGTOLFEDNALAVLDNG 120  
 Db 61 ELTYLPTNASLFLDIOIEYQGVYVLAHQVQVPLQRLRYNGTOLFEDNALAVLDNG 120  
 QY 121 DPLNNTPTVGASPGGLRELOLRSLEIKGGVLIQNPOLCYODTILMKDIFHKNNOLA 180  
 Db 121 DPLNNTPTVGASPGGLRELOLRSLEIKGGVLIQNPOLCYODTILMKDIFHKNNOLA 180  
 QY 181 LTLIDNRSRACHPCSPMKGSRGWSESDCOSLRTVCAGGACARCKPLPTDCHEQC 240  
 Db 181 LTLIDNRSRACHPCSPMKGSRGWSESDCOSLRTVCAGGACARCKPLPTDCHEQC 240  
 QY 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVTYNTDTESMPNPRGRTTFCASCTYAC 300  
 Db 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVTYNTDTESMPNPRGRTTFCASCTYAC 300  
 QY 301 YNLTSDVSGCTLVCPHANOEVTAEDGTORCEKSKPCARVCYGLGMEHLREVRATVSAN 360  
 Db 301 YNLTSDVSGCTLVCPHANOEVTAEDGTORCEKSKPCARVCYGLGMEHLREVRATVSAN 360  
 QY 361 IOEFACCKRIFGSLAFLEPSFDGDPASNTAPLQPEOLOVFELEETIGLYISAMPDSLP 420  
 Db 361 IOEFACCKRIFGSLAFLEPSFDGDPASNTAPLQPEOLOVFELEETIGLYISAMPDSLP 420  
 QY 421 DLSVQNTQVIRGRILHNAYSLTLOGIGISWGLRSLRELSGLALIHNNHLCFVHTV 480  
 Db 421 DLSVQNTQVIRGRILHNAYSLTLOGIGISWGLRSLRELSGLALIHNNHLCFVHTV 480  
 QY 481 PMDOLFRNPHQALLHTANPEDECEVGEGLACHQOLCARGHCMGPGPTQVCNSQFLRGQEC 540  
 Db 481 PMDOLFRNPHQALLHTANPEDECEVGEGLACHQOLCARGHCMGPGPTQVCNSQFLRGQEC 540  
 QY 541 VEECRVLOGLPREYVNAHRLCPCHRECOPONGSVTCFGEADQCVACAHYKDPPECVARC 600  
 Db 541 VEECRVLOGLPREYVNAHRLCPCHRECOPONGSVTCFGEADQCVACAHYKDPPECVARC 600  
 QY 601 PSQVPRDLSYMPIMKFPDEEGACQPCPINCTHSCVDLDKGCFAEQRASPLTSLISAVYG 660  
 Db 601 PSQVPRDLSYMPIMKFPDEEGACQPCPINCTHSCVDLDKGCFAEQRASPLTSLISAVYG 660  
 QY 654 ----- 653  
 Db 661 ILLVVLGVVFGILLIRROOKIRKYMRLLOETELVEPLTPSGAMPNOQMRLKETEL 720  
 QY 654 ----- 653  
 Db 721 RKVKVLGSGAFGVYKGIWIPDGENVKIPVAIKVIRENTSPKANKELDEAYMAGVGSF 780  
 QY 654 ----- 653  
 Db 781 YVSRLLGICLTSTVOLVTOQMPYGCLLDHVRENRRGLSGODLLNMCQIAKMSYLEDR 840  
 QY 654 ----- 653



Db 1081 AGSDVDFDGLMGAKGLQSLPTHTDPSLQRTSEDPVPLPSEDTGCVAPLTCSPQPEYV 1140  
QY 805 NOPDVRPOPSPRPGPLPAARPAAGATLERPKTSLPGKNGYKVDVAFAGAVENPEYLTPQ 864  
Db 1141 NOPDVRPOPSPRPGPLPAARPAAGATLERPKTSLPGKNGYKVDVAFAGAVENPEYLTPQ 1200  
QY 865 GGAAPQHPPPAPSPAFDNLVYMDODPPERGAPSPSTFKGPTAENPEYLGLDVPV 919  
Db 1201 GGAAPQHPPPAPSPAFDNLVYMDODPPERGAPSPSTFKGPTAENPEYLGLDVPV 1255

RESULT 7  
US-08-466-680B-68  
; Sequence 68, Application US/08466680B  
; Patent No. 6075122  
; GENERAL INFORMATION:  
; APPLICANT: Cheever, Martin A.  
; APPLICANT: Disis, Mary L.  
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN  
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed and Berry LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: US  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/466,680B  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharkey, Richard G.  
; REGISTRATION NUMBER: 32,629  
; REFERENCE/DOCKET NUMBER: 920010.448C4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 68:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1255 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-466-680B-68

Query Match 96.3%; Score 4891; DB 3; Length 1255;  
Best Local Similarity 73.1%; Pred. No. 0;  
Matches 918; Conservative 0; Mismatches 1; Indels 336; Gaps 1;

QY 1 MELAALCRWGLLALLLPFGAASVCTGTDMKRLPASPEETHLMDLRLHYOGCOVVOGNTL 60  
Db 1 MELAALCRWGLLALLLPFGAASVCTGTDMKRLPASPEETHLMDLRLHYOGCOVVOGNTL 60  
QY 61 ELTYLTPNASLSFLQDIOEVGYVLIHNOVOYVPLQRLRIVRGSTOLEFENYALAVYDNG 120  
Db 61 ELTYLTPNASLSFLQDIOEVGYVLIHNOVOYVPLQRLRIVRGSTOLEFENYALAVYDNG 120  
QY 121 DPLNNTTPVYGASPGGLRELQSLTEILKGVLIQNRPOLCYODTILMKDIFKKNOLA 180  
Db 121 DPLNNTTPVYGASPGGLRELQSLTEILKGVLIQNRPOLCYODTILMKDIFKKNOLA 180  
QY 181 LTLIDTNRSRACHPCSPMGKSGSRMGSSSEDCQSLTRTVAGGACARKGPLPTDCCHEQC 240  
Db 181 LTLIDTNRSRACHPCSPMGKSGSRMGSSSEDCQSLTRTVAGGACARKGPLPTDCCHEQC 240  
QY 241 AAGCTGPKHSDCLACLFHNHSGICELHCPALVYNTDTFESMPNDEGRYTFGACVYACP 300

Db 241 AAGCTGPKHSDCLACLFHNHSGICELHCPALVYNTDTFESMPNDEGRYTFGACVYACP 300  
QY 301 YNLTSDVGSCTIVCPILHNDVTAEDGTORCEKSKPCARCYGLGHEHLREVAVTSAN 360  
Db 301 YNLTSDVGSCTIVCPILHNDVTAEDGTORCEKSKPCARCYGLGHEHLREVAVTSAN 360  
QY 361 IOEFAGCKITFSLAFIPESFDGPASNTAPLOEQQLQVFETLEITGYLITSMPDLSL 420  
Db 361 IOEFAGCKITFSLAFIPESFDGPASNTAPLOEQQLQVFETLEITGYLITSMPDLSL 420  
QY 421 DLSVFQNLQVIRGRILHNGAVSLTLQGLISLWLSRLRELGSGLALIHNTHTLCFVHTV 480  
Db 421 DLSVFQNLQVIRGRILHNGAVSLTLQGLISLWLSRLRELGSGLALIHNTHTLCFVHTV 480  
QY 481 PMDQLFNPHQALLHTANRPDEBCEVGEGLACHQLCARHCWGPPTQCVCNCSQFLRGQEC 540  
Db 481 PMDQLFNPHQALLHTANRPDEBCEVGEGLACHQLCARHCWGPPTQCVCNCSQFLRGQEC 540  
QY 541 VEECRVILQGLPREVYVNAHCLPCRPBQOPONGSVTCGPEADOCVACAHKKDPFCVARC 600  
Db 541 VEECRVILQGLPREVYVNAHCLPCRPBQOPONGSVTCGPEADOCVACAHKKDPFCVARC 600  
QY 601 PSQVCPDLSYMPIMKFPDEBACOPCPINCTHSCVDLDDKCPAPORASPLTS----- 653  
Db 601 PSQVCPDLSYMPIMKFPDEBACOPCPINCTHSCVDLDDKCPAPORASPLTS----- 653  
QY 654 ----- 653  
Db 654 ----- 653  
QY 661 ILVYVLGVVFGILIKRRQOKIRKRYTMRLLOETELVEPLTPSGAMPNQAOMRIKTEL 720  
Db 661 ILVYVLGVVFGILIKRRQOKIRKRYTMRLLOETELVEPLTPSGAMPNQAOMRIKTEL 720  
QY 654 ----- 653  
Db 654 ----- 653  
QY 721 RKVAVLGSAGFTVYKGIWIPDGENVKI PVAIKVRENTSPKANKETIDEAYVAGVGP 780  
Db 721 RKVAVLGSAGFTVYKGIWIPDGENVKI PVAIKVRENTSPKANKETIDEAYVAGVGP 780  
QY 654 ----- 653  
Db 654 ----- 653  
QY 841 LVHRDLAARNVLYKSPNHVKITDFGLARLIDIDETEXHADGKVPKIMMALESLRRFT 900  
Db 841 LVHRDLAARNVLYKSPNHVKITDFGLARLIDIDETEXHADGKVPKIMMALESLRRFT 900  
QY 654 ----- 653  
Db 654 ----- 653  
QY 901 HQSDVMSYGVYVWELMTFGAKRYDGLIPAREIPLLEKGERLPOPICTIDVYIMVCM 960  
Db 901 HQSDVMSYGVYVWELMTFGAKRYDGLIPAREIPLLEKGERLPOPICTIDVYIMVCM 960  
QY 654 ----- 653  
Db 654 ----- 653  
QY 961 IDSECRPFRELVSEFSRMAADPQRFVYIQNEDLGPASPLDSTFRSLLEDDEDDGLVDA 1020  
Db 961 IDSECRPFRELVSEFSRMAADPQRFVYIQNEDLGPASPLDSTFRSLLEDDEDDGLVDA 1020  
QY 685 EBYLVPOOGFECPPDAPAGAGVYHHRSSSTRSGGDLTLGLEPSEEA PRSLAPSEG 744  
Db 1021 EBYLVPOOGFECPPDAPAGAGVYHHRSSSTRSGGDLTLGLEPSEEA PRSLAPSEG 744  
QY 745 AGSVYFPGDLMGAKGLQSLPTHTDPSLQRTSEDPVPLPSEDTGCVAPLTCSPQPEYV 804  
Db 1081 AGSVYFPGDLMGAKGLQSLPTHTDPSLQRTSEDPVPLPSEDTGCVAPLTCSPQPEYV 804  
QY 805 NOPDVRPOPSPRPGPLPAARPAAGATLERPKTSLPGKNGYKVDVAFAGAVENPEYLTPQ 864  
Db 1141 NOPDVRPOPSPRPGPLPAARPAAGATLERPKTSLPGKNGYKVDVAFAGAVENPEYLTPQ 864  
QY 865 GGAAPQHPPPAPSPAFDNLVYMDODPPERGAPSPSTFKGPTAENPEYLGLDVPV 919  
Db 1201 GGAAPQHPPPAPSPAFDNLVYMDODPPERGAPSPSTFKGPTAENPEYLGLDVPV 919

RESULT 8  
US-08-484-438-8  
; Sequence 8, Application US/08484438  
; Patent No. 5811098  
; Patent No. 5811098 5780031

GENERAL INFORMATION:  
APPLICANT: Plozman, Gregory D.  
APPLICANT: Culouscou, Jean-Michel  
APPLICANT: Shoyab, Mohammed  
APPLICANT: Stegall, Clay B.  
APPLICANT: Helistr m, Ingegerd  
APPLICANT: Helistr m, Karl E.  
TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,438  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/323,442  
FILING DATE: 14-OCT-1994  
APPLICATION NUMBER: US 08/150,704  
FILING DATE: 10-NOV-1993  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/981,165  
FILING DATE: 24-NOV-1992  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-230  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNITE  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1255 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-484-438-8

Query Match 94.9% Score 4818; DB 2; Length 1255;  
Best Local Similarity 72.6%; Pred. No. 0;  
Matches 912; Conservative 0; Mismatches 6; Indels 338; Gaps 3;

QY 1 MELAALCRWGLLLALLPGAASTOYCTGTDMKLRLPASBETHLMDLRHLYOGCOVVGNL 60  
DB 1 MELAALCRWGLLLALLPGAASTOYCTGTDMKLRLPASBETHLMDLRHLYOGCOVVGNL 60  
QY 61 ELTYPTNASLSFLDIOGVGVLAHNOVQVPLQRLRIYRGTOLEFDNTALAVLDNG 120  
DB 61 ELTYPTNASLSFLDIOGVGVLAHNOVQVPLQRLRIYRGTOLEFDNTALAVLDNG 120  
QY 121 DELNNTPTVTGASPGGLRELOLRSLTEILKGVLIORNPOLCYODTILMKDIFHKNNOLA 180  
DB 121 DELNNTPTVTGASPGGLRELOLRSLTEILKGVLIORNPOLCYODTILMKDIFHKNNOLA 180  
QY 181 LTLIDTNRSRACHPCSPMKSGSRGSESESDCQSLRTVVCAGACACAKGPLPTDCHEOC 240  
DB 181 LTLIDTNRSRACHPCSPMKSGSRGSESESDCQSLRTVVCAGACACAKGPLPTDCHEOC 240  
QY 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDTFESMPNPGRYTFGASCVTACP 300

DB 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDTFESMPNPGRYTFGASCVTACP 300  
QY 301 YNTLSTDVSCITVYGLHNOEYTAEDGTORCEKSCPCARVCGIGMEHLREVRATSN 360  
DB 301 YNTLSTDVSCITVYGLHNOEYTAEDGTORCEKSCPCARVCGIGMEHLREVRATSN 360  
QY 361 IOEFACCKRIFGSLAFLEPSFDGDPASNTAPLOPEOLOVFETLEETGYLYISAMPDLP 420  
DB 361 IOEFACCKRIFGSLAFLEPSFDGDPASNTAPLOPEOLOVFETLEETGYLYISAMPDLP 420  
QY 421 DLSVFONLOYIRGRILHNCAYSITLLOGIGISWLGRLSLELSGGLAIHHNTHLCVHTV 480  
DB 421 DLSVFONLOYIRGRILHNCAYSITLLOGIGISWLGRLSLELSGGLAIHHNTHLCVHTV 480  
QY 481 PMDQLFRNPHOALHTANRDECEVEGLACHOLGARGHMGVPTQCYNCSOFLRGOC 540  
DB 481 PMDQLFRNPHOALHTANRDECEVEGLACHOLGARGHMGVPTQCYNCSOFLRGOC 540  
QY 541 VEECRVLOGLPREYVNAHRLCPHPECOPQNGSVTCFGEADQCVACAHYKDPPECVARC 600  
DB 541 VEECRVLOGLPREYVNAHRLCPHPECOPQNGSVTCFGEADQCVACAHYKDPPECVARC 600  
QY 601 PSGVPELSTYMPITWKPDEEGACQPCPINCTHSCVDLDDKGPABORASPLTS----- 653  
DB 601 PSGVPELSTYMPITWKPDEEGACQPCPINCTHSCVDLDDKGPABORASPLTSIYSAVVG 660  
QY 654 ----- 653  
DB 654 ILLVVVLGVVFGILIKRRQOKIRKTYMRLOETELVEPLPESGAMPNOAMRIKETEL 720  
QY 654 ----- 653  
DB 721 RKVKVLSGAFGVYKGIWIPGENYKIPAIKVLRENTSPRANKRIIDEAVYMGVGS 780  
QY 654 ----- 653  
DB 781 YVSRLLIGLSTVQLVTLQMPYGLLDHVRENRGLGSQDLNMCQIAKGSYLEVDR 840  
QY 654 ----- 653  
DB 841 LVHRDLAARNLVKSPNHVKITDFGLARLIDDEYEHADGGKVPITKMMALSLARRPT 900  
QY 654 ----- 653  
DB 901 HQSDVMSGVTVWELMTFGAKPYDGIPIAREIPDLLEKGRPLPQPICTIDVYIMVYKWM 960  
QY 654 ----- QNEDLGASFLDSTFFYSLLEDDDMGDLVA 684  
DB 961 IDSECRPRRELIVSEFSRMARDPQRFVVTQNEDLGASPLDSTFFYSLLEDDMGDLVA 1020  
QY 685 EEYLVPOQGFECDDPAPGAGWVHHRSSSTRSGGDLTLGLPESEEPAPSPPLAPSEG 744  
DB 1021 EEYLVPOQGFECDDPAPGAGWVHHRSSSTRSGGDLTLGLPESEEPAPSPPLAPSEG 1080  
QY 745 ACSDFVFDGLGMAKAGLOSPLTHDPSPLQRYSEDPVLPBSETDGYVAPLTCSPQEXY 804  
DB 1081 ACSDFVFDGLGMAKAGLOSPLTHDPSPLQRYSEDPVLPBSETDGYVAPLTCSPQEXY 1140  
QY 805 NOPDVAPRPPSPREGPLPAARAGATLERKTLSPKKNVYVDVAFGAVENREYLTQ 864  
DB 1141 NOPDVAPRPPSPREGPLPAARAGATLERKTLSPKKNVYVDVAFGAVENREYLTQ 1200  
QY 865 GGAAPQPHPPAFSPAFDNLVYWDOPPRGAPSPSTFKGPT-AEPEYIGLDVVP 919  
DB 1201 GGAAPQPHPPAFSPAFDNLVYWDOPPRGAPSPSTFKGPTAENPEY-GLDVPV 1255

RESULT 9  
US-09-146-283-4  
; Sequence 4, Application US/09146283  
; Patent No. 5976546  
; GENERAL INFORMATION:

	APPLICANT:	Laus, Reinert
	APPLICANT:	Ruegg, Curtis L.
	APPLICANT:	Wu, Hongyu
	TITLE OF INVENTION:	Immunostimulatory Compositions
	NUMBER OF SEQUENCES:	10
	CORRESPONDENCE ADDRESS:	
	ADDRESSEE:	Dehlinger & Associates
	STREET:	350 Cambridge Ave. Suite 250
	CITY:	Palo Alto
	STATE:	CA
	COUNTRY:	USA
	ZIP:	94306
	COMPUTER READABLE FORM:	
	MEDIUM TYPE:	Floppy disk
	COMPUTER:	IBM PC compatible
	OPERATING SYSTEM:	PC-DOS/MS-DOS
	SOFTWARE:	Patentin Release #1.0, Version #1.25
	CURRENT APPLICATION DATA:	
	APPLICATION NUMBER:	US/09/146, 283
	FILING DATE:	03-SEPT-1998
	CLASSIFICATION:	536
	ATTORNEY/AGENT INFORMATION:	
	NAME:	Judge, Linda R.
	REGISTRATION NUMBER:	42,702
	REFERENCE/DOCKET NUMBER:	7636-0010.21
	TELECOMMUNICATION INFORMATION:	
	TELEPHONE:	650-324-0880
	TELEFAX:	650-324-0960
	INFORMATION FOR SEQ ID NO:	4:
	SEQUENCE CHARACTERISTICS:	
	LENGTH:	782 amino acids
	TYPE:	amino acid
	TOPOLOGY:	linear
	MOLECULE TYPE:	protein
	HYPOTHETICAL:	NO
	ORIGINAL SOURCE:	
	ORGANISM:	homo sapiens
	INDIVIDUAL ISOLATE:	GM-CSF-Her-2 fusion protein; Fig. 8
	US-09-146-283-4	

	Nucleary Match	71.5%	Score 3632;	DB 2;	Length 782;
	Fast Local Similarity	98.7%;	Pred. No. 1.3e-271;		
	Matches 659;	Conservative 0;	Mismatches 5;	Indels 4;	Gaps 2.

	1	MELALICRMGLLALLLP	GAAS	TVCTGTD	MDKLR	LPA	SPE	THLDMLR	HL	YOC	GVQ	GNL	60																																													
Dd	1	MELALICRMGLLALLP	PGAS	TVCTGTD	MDKLR	LPA	SPE	THLDMLR	HL	YOC	GVQ	GNL	60																																													
Qy	61	E	LTY	P	TNAS	L	S	F	L	D	I	O	E	V	O	G	V	L	I	A	H	N	O	V	R	O	V	P	L	O	R	L	R	I	V	R	G	T	O	L	E	D	N	T	A	L	A	V	D	N	G	120						
Dd	61	E	LTY	P	TNAS	L	S	F	L	D	I	O	E	V	O	G	V	L	I	A	H	N	O	V	R	O	V	P	L	O	R	L	R	I	V	R	G	T	O	L	E	D	N	T	A	L	A	V	D	N	G	120						
Qy	121	D	P	L	N	T	T	V	T	A	S	P	G	L	E	L	D	R	S	T	E	L	K	G	V	L	L	O	R	N	P	O	L	C	Y	O	D	T	I	L	M	K	D	I	F	H	K	N	N	O	L	A	180					
Dd	121	D	P	L	N	T	T	V	T	A	S	P	G	L	E	L	D	R	S	T	E	L	K	G	V	L	L	O	R	N	P	O	L	C	Y	O	D	T	I	L	M	K	D	I	F	H	K	N	N	O	L	A	180					
Qy	181	L	T	L	I	D	T	N	S	R	A	C	H	P	S	P	M	C	K	S	R	C	W	G	S	S	E	S	E	D	C	O	S	L	T	R	V	C	A	G	A	C	A	R	K	G	P	L	T	D	C	H	E	O	C	240		
Dd	181	L	T	L	I	D	T	N	S	R	A	C	H	P	S	P	M	C	K	S	R	C	W	G	S	S	E	S	E	D	C	O	S	L	T	R	V	C	A	G	A	C	A	R	K	G	P	L	T	D	C	H	E	O	C	240		
Qy	241	A	A	G	C	T	P	R	H	S	D	C	L	A	C	L	A	H	N	S	G	I	C	E	L	A	C	P	A	L	V	T	T	M	T	D	F	F	S	M	P	N	D	E	G	R	T	T	F	G	A	S	C	V	T	A	C	300
Dd	241	A	A	G	C	T	P	R	H	S	D	C	L	A	C	L	A	H	N	S	G	I	C	E	L	A	C	P	A	L	V	T	T	M	T	D	F	F	S	M	P	N	D	E	G	R	T	T	F	G	A	S	C	V	T	A	C	300
Qy	301	Y	N	T	S	T	D	V	G	S	C	T	I	V	C	P	L	N	O	E	V	T	A	D	G	O	R	E	K	S	K	P	C	A	R	V	C	Y	G	I	G	M	E	L	R	V	R	A	V	T	S	N	360					
Dd	301	Y</																																																								

```

Db      421  DLVSFQNIQYINGRIILHNATSLYTIQIGISTWLGRLSRLRELSGIALIHNNHILCVHY 480
QY      481  PWDOLFRRPHQALHTANRPEDECVGEGLACHQLCARHGCHWGPPTQVCNCSQFLRGDEC 540
Db      481  PWDOLFRRPHQALHTANRPEDECVGEGLACHQLCARHGCHWGPPTQVCNCSQFLRGDEC 540
QY      541  VEECHVIOGLPREYVYNNARHCLPCHPECCQPNQNSVYTCFGEADQCVACAHYKDPFVCAR 600
Db      541  VEECHVIOGLPREYVYNNARHCLPCHPECCQPNQNSVYTCFGEADQCVACAHYKDPFVCAR 600
QY      601  PSGVPRDLSTYMWKFPDEEGACQCPINCTHSCVDLDLDKGPAPGROASPLTSQNEIDLAP 660
Db      601  PSGVPRDLSTYMWKFPDEEGACQCPINCTHSCVDLDLDKGPAPGROASPLTSLE---AP 657
QY      661  A-SPLDST 667
Db      658  ARSPSPST 665

RESULT 10
US-08-579-823A-4
: Sequence 4, Application US/08579823A
: Patent No. 6080409
:
: GENERAL INFORMATION:
: APPLICANT: Laus, Reinert
: APPLICANT: Ruegg, Curtis L.
: TITLE OF INVENTION: Immunostimulatory Composition and Method
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Dehlinger & Associates
: STREET: 350 Cambridge Ave. Suite 250
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94306
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/579,823A
: FILING DATE: 03-DEC-1998
: CLASSIFICATION: 536
: ATTORNEY/AGENT INFORMATION:
: NAME: Judge, Linda R.
: REGISTRATION NUMBER: 42,702
: REFERENCE/DOCKET NUMBER: 7636-0010
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 650-324-0960
: TELEFAX: 650-324-0960
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 782 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: HYPOTHETICAL: NO
: ORIGINAL SOURCE:
: ORGANISM: homo sapiens
: INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8
US-08-579-823A-4

Query Match      71.5%; Score 3632; DB 3; Length 782;
Best Local Similarity 98.7%; Pred. No. 1,3e-27;
Matches 659; Conservative 0; Mismatches 5; Indels 4; Gaps 2,

```



RESULT 12  
US-08-422-108-1  
Sequence 1, Application US/08422108  
Patent No. 6015567  
GENERAL INFORMATION:  
APPLICANT: Huddziak, Robert M.  
APPLICANT: Shepard, H. Michael  
APPLICANT: Ullrich, Axel  
TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/422,108  
FILING DATE: 14-Apr-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/355460  
FILING DATE: 13-DEC-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/048346  
FILING DATE: 15-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/354319  
FILING DATE: 19-MAY-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M  
REGISTRATION NUMBER: 00,000  
REFERENCE/DOCKET NUMBER: 554C2D2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 624 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
US-08-422-108-1

Query Match 68.4%; Score 3473; DB 3; Length 624;  
Best Local Similarity 99.8%; Pred. No. 1.7e-259;  
Matches 623; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 22 STGVCTGTMTKRLPASPETHLDMRLHYOGGVVGNLFTYLPNTMSLFDIOEYQ 81  
DB 1 STGVCTGTMTKRLPASPETHLDMRLHYOGGVVGNLFTYLPNTMSLFDIOEYQ 60  
QY 82 GYVLIAHNOVROVPLQRLRIYRGTLFEDNYALAVLDNGDPLNNTTPTVGASPGGLREIQ 141  
DB 61 GYVLIAHNOVROVPLQRLRIYRGTLFEDNYALAVLDNGDPLNNTTPTVGASPGGLREIQ 120  
QY 142 LRSLEIILKGVLIQRNPQLCYQDPTILMKDIFHKNNQALATLIDNRSACHPSCPMCKG 201  
DB 121 LRSLEIILKGVLIQRNPQLCYQDPTILMKDIFHKNNQALATLIDNRSACHPSCPMCKG 180  
QY 202 SRMGESSSDCSLRTFYCAGGACRCKGPLPTDCCHBEOCAAGCTGPKHSDCLACLHFHNS 261  
DB 181 SRMGESSSDCSLRTFYCAGGACRCKGPLPTDCCHBEOCAAGCTGPKHSDCLACLHFHNS 240  
QY 262 GICELHCPALVYNTDFESMNPREGRTFGASCVTACPYNYLSIDVGSCTLVCEPLHNOE 321  
DB 321 GICELHCPALVYNTDFESMNPREGRTFGASCVTACPYNYLSIDVGSCTLVCEPLHNOE 300

DB 241 GICELHCPALVYNTDFESMNPREGRTFGASCVTACPYNYLSIDVGSCTLVCEPLHNOE 300  
QY 322 VTAEDGTORCEKSKPCARVCYGLGMEHLREYRAVTSANIOFAGCKITFGSLAFPESEF 381  
DB 301 VTAEDGTORCEKSKPCARVCYGLGMEHLREYRAVTSANIOFAGCKITFGSLAFPESEF 360  
QY 382 DDDPASNTAPLOPEOLQVETLEETIGYLYISAMPDLPDLVSFONLQYIRGRILHNGAY 441  
DB 361 DDDPASNTAPLOPEOLQVETLEETIGYLYISAMPDLPDLVSFONLQYIRGRILHNGAY 420  
QY 442 SLTLQGLIGISWLGRLSRLBGLSGLALIHNTLHCFVHTVPMQDLFENPHQALLHTANPE 501  
DB 421 SLTLQGLIGISWLGRLSRLBGLSGLALIHNTLHCFVHTVPMQDLFENPHQALLHTANPE 480  
QY 502 DECVEGGLACRQLCARGHGWGPGPTQCVVCSQFLRQGEVEECRYLQGLPREYVNAHRL 561  
DB 481 DECVEGGLACRQLCARGHGWGPGPTQCVVCSQFLRQGEVEECRYLQGLPREYVNAHRL 540  
QY 562 PCHECQPOGNSVTCGPPADOCVACAHYKDDPPFCVARGCPGKPDLSYMPIMKPPDEEG 621  
DB 541 PCHECQPOGNSVTCGPPADOCVACAHYKDDPPFCVARGCPGKPDLSYMPIMKPPDEEG 600  
QY 622 ACQPCPINCTHSCVDDLDKGCPEAE 645  
DB 601 ACQPCPINCTHSCVDDLDKGCPEAE 624

RESULT 13  
US-08-422-734-1  
Sequence 1, Application US/08422734  
Patent No. 633169  
GENERAL INFORMATION:  
APPLICANT: Huddziak, Robert M.  
APPLICANT: Shepard, H. Michael  
APPLICANT: Ullrich, Axel  
TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/422,734  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/422108  
FILING DATE: 14-Apr-1995  
APPLICATION NUMBER: 08/354460  
FILING DATE: 13-DEC-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/048346  
FILING DATE: 15-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/354319  
FILING DATE: 19-MAY-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M  
REGISTRATION NUMBER: 00,000  
REFERENCE/DOCKET NUMBER: 554C2D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 1:



SEQUENCE CHARACTERISTICS:  
 LENGTH: 624 amino acids  
 TYPE: Amino Acid  
 TOPOLOGY: Linear  
 US-08-422-734-1

Query Match 68.4%; Score 3473; DB 4; Length 624;  
 Best Local Similarity 99.8%; Pred. No. 1.7e-259;  
 Matches 623; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

22 STGVCTGDKMLRPAPEHMLDMLRHLVGGCOVGNLFTLYPTNASLFTLDDIOEQ 81  
 1 STGVCTGDKMLRPAPEHMLDMLRHLVGGCOVGNLFTLYPTNASLFTLDDIOEQ 60  
 82 GYVLIANOVROYPLQRLIRYRGTOLEFEDNYALAVLDNGDPRLNNTTPTVGSPEGLREIQ 141  
 61 GYVLIANOVROYPLQRLIRYRGTOLEFEDNYALAVLDNGDPRLNNTTPTVGSPEGLREIQ 120  
 142 LRSITLILKGGVLIQRPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPMCKG 201  
 121 LRSITLILKGGVLIQRPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPMCKG 180  
 202 SRCGSESEDCQSLTRTVCAAGGCAKCKPLPTDCHEQCAAGCTGPKHSDLAFLHNS 261  
 181 SRCGSESEDCQSLTRTVCAAGGCAKCKPLPTDCHEQCAAGCTGPKHSDLAFLHNS 240  
 262 GICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGSCTLVCPHNOE 321  
 241 GICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGSCTLVCPHNOE 300  
 322 VTADGTQRECKSKPCARVCYGLGMEHLREVAVTSANIOEFAGCKKIFGSLAFLESEF 381  
 301 VTADGTQRECKSKPCARVCYGLGMEHLREVAVTSANIOEFAGCKKIFGSLAFLESEF 360  
 382 DGDPASTARLPOLQVFELEETITGLYISAMPDLSPLDSVONOLQVIRGLHNCAY 441  
 361 DGDPASTARLPOLQVFELEETITGLYISAMPDLSPLDSVONOLQVIRGLHNCAY 420  
 442 SLTLOGIGISWLGRLSRELSSGLALIHNTLHLCFVHTVPMDQLFRPHQALHTANPE 501  
 421 SLTLOGIGISWLGRLSRELSSGLALIHNTLHLCFVHTVPMDQLFRPHQALHTANPE 480  
 502 DECVBELACHQLCARHCWPGPTQCVNCSQFLRGQECVECHVLOGLPREYVNAHCL 561  
 481 DECVBELACHQLCARHCWPGPTQCVNCSQFLRGQECVECHVLOGLPREYVNAHCL 540  
 562 PCHEGCPONGSVTCFGEADQCAAHYKDRPCVACRPSGVKPRDLSYMPIMKFPDEEG 621  
 541 PCHEGCPONGSVTCFGEADQCAAHYKDRPCVACRPSGVKPRDLSYMPIMKFPDEEG 600  
 622 ACQPCPINCTHSCVDLDDKGPAP 645  
 601 ACQPCPINCTHSCVDLDDKGPAP 624

RESULT 14  
 US-09-630-155-2  
 ; Sequence 2, Application US/09630155  
 ; Patent No. 641430

GENERAL INFORMATION:  
 APPLICANT: Doherty, Joni Kristin and Gail M. Clinton  
 TITLE OF INVENTION: HER-2 BINDING ANTAGONISTS  
 NUMBER OF SEQUENCES: 9  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: DAVIS WRIGHT TREMAINE LLP  
 STREET: 1501 Fourth Avenue, 2600 Century Square  
 CITY: Seattle  
 STATE: Washington  
 COUNTRY: U.S.A.  
 ZIP: 98101  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: PC compatible

OPERATING SYSTEM: Windows95  
 SOFTWARE: Word  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: us/09/630,155  
 FILING DATE: 16-Jan-2001  
 CLASSIFICATION: <Unknown>  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Davison, Barry L.  
 REGISTRATION NUMBER: 47,309  
 REFERENCE/DOCKET NUMBER: 49321-10  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 206 628-7621  
 TELEFAX: 206 628-7699  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 419  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: unknown  
 MOLECULE TYPE: polypeptide  
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
 US-09-630-155-2

Query Match 37.0%; Score 1878; DB 4; Length 419;  
 Best Local Similarity 83.0%; Pred. No. 8.3e-137;  
 Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

1 MELALCRWGLLALPPGASTQVCTGDKMLRPAPEHMLDMLRHLVGGCOVGNL 60  
 1 MELALCRWGLLALPPGASTQVCTGDKMLRPAPEHMLDMLRHLVGGCOVGNL 60  
 61 ELTYLPTNASLFTLDDIOEQGVYLIANOVROYPLQRLIRYRGTOLEFEDNYALAVLDNG 120  
 61 ELTYLPTNASLFTLDDIOEQGVYLIANOVROYPLQRLIRYRGTOLEFEDNYALAVLDNG 120  
 121 DPLNNTPTVYGASGGLRELDLRSITLILKGGVLIQRPOLCYODTILMKDIFHKNNOLA 180  
 121 DPLNNTPTVYGASGGLRELDLRSITLILKGGVLIQRPOLCYODTILMKDIFHKNNOLA 180  
 181 LTLIDTNSRACHPCSPMCKSKRCMGESSEDCQSLTRTVCAAGGCAKCKPLPTDCHEOC 240  
 181 LTLIDTNSRACHPCSPMCKSKRCMGESSEDCQSLTRTVCAAGGCAKCKPLPTDCHEOC 240  
 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESMPNPEGRTFGASCVTACP 300  
 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESMPNPEGRTFGASCVTACP 300  
 301 YNYLSTDVGSCTLVCPHNOEVTADGTQRECKSKPCAR----GTHSLRPAAVPPV 355  
 301 YNYLSTDVGSCTLVCPHNOEVTADGTQRECKSKPCAR----GTHSLRPAAVPPV 355  
 361 IOEFAGCKKIRGSLAFLESEFADGPASNT---APLOPQOLQVFELEETITGLYISAMPD 417  
 361 IOEFAGCKKIRGSLAFLESEFADGPASNT---APLOPQOLQVFELEETITGLYISAMPD 417  
 418 SLPDLVSFONLQVIRG 433  
 406 --PDAAHVAVNLNRYEG 419

RESULT 15  
 US-08-484-438-7  
 ; Sequence 7, Application US/08484438  
 ; Patent No. 5811098  
 ; Patent No. 5811098 5780031

GENERAL INFORMATION:  
 APPLICANT: Plowman, Gregory D.  
 APPLICANT: Culouscou, Jean-Michel  
 APPLICANT: Shoyab, Mohammed  
 APPLICANT: Siegall, Clay B.  
 APPLICANT: Hellstr m, Ingegerd  
 APPLICANT: Hellstr m, Karl E.  
 TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE

```

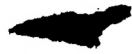
1 NUMBER OF SEQUENCES: 42
2
3 CORRESPONDENCE ADDRESS:
4 ADDRESSEE: Pennie & Edmonds
5 STREET: 115 Avenue of the Americas
6 CITY: New York
7 STATE: New York
8 COUNTRY: U.S.A.
9 ZIP: 10036-2711
10
11 COMPUTER READABLE FORM:
12 MEDIUM TYPE: Floppy disk
13 COMPUTER: IBM PC compatible
14 OPERATING SYSTEM: PC-DOS/MS-DOS
15 SOFTWARE: Patent Release #1.0, Version #1.25
16
17 CURRENT APPLICATION DATA:
18 APPLICATION NUMBER: US/08/484,438
19 FILING DATE: 07-JUN-1995
20 CLASSIFICATION: 530
21 PRIOR APPLICATION DATA:
22 APPLICATION NUMBER: 08/323,442
23 FILING DATE: 14-OCT-1994
24 APPLICATION NUMBER: US 08/150,704
25 FILING DATE: 10-NOV-1993
26 CLASSIFICATION: 530
27 PRIOR APPLICATION DATA:
28 APPLICATION NUMBER: US 07/981,165
29 FILING DATE: 24-NOV-1992
30 CLASSIFICATION: 530
31 ATTORNEY/AGENT INFORMATION:
32 NAME: Mistrock, S. Leslie
33 REGISTRATION NUMBER: 18,872
34 REFERENCE/DOCKET NUMBER: 5624-230
35 TELECOMMUNICATION INFORMATION:
36 TELEPHONE: (212) 790-9090
37 TELEFAX: (212) 869-8864/9741
38 TELEX: 66141 PENNIE
39 INFORMATION FOR SEQ ID NO: 7:
40 SEQUENCE CHARACTERISTICS:
41 LENGTH: 1210 amino acids
42 TYPE: amino acid
43 STRANDEDNESS: unknown
44 TOPOLOGY: unknown
45 MOLECULE TYPE: protein
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
95
```

```

Db      363  TSISGDHLHPVAARGSFTHTPPLDPOEDLDITKYKEITGELLIDAMENRTDLHAFEN  422
Qy      428  LQVIRGRILNHGASLTLTQAGISIMWLSRSLRELGLALIHNNTHLCVHTVPMDQLEF  487
Db      423  LEIRINGRTKQGOESLAVNSINTSISGLRSLKEISDGDVYISGNKNLVCYANTINMKKLEF  482
Qy      488  NPHQALLHTANRPDEBCEVSGGLACHOLACAGHCHMGEGFPQCVNCSQFLGQCEVECRVL  547
Db      483  TSGQRTKIISIRGENSCKATGOYCHALCSPEGGGEPBPCVSCRNVSGRRCVCKLL  542
Qy      548  QGLPREVYNABHCPHCEPCOPONGSVTCGPEPADCCVACAHKDPDFVACPSGVKPD  607
Db      543  EGEPREREVENECQCHPREDLPQAMNITCTGRPDNCTICQAHIDDPHCVKCPAGVME  602
Qy      608  LSYMPKFPDEBEGACOPCPPICTHSCVDLDDKGCFA-----  644
Db      603  NNTL-VMKYADAGHVCHLCPNCTYGGCTGGLGCEPCTNGKPIPSIATGMVGAALLLVYA  661
Qy      645  -----EORASLTQNE-----  656
Db      662  LGIGLFMRRIIVAKRILRLRLODERELVEPLTSGSAPNALLILKETEFKIKVLSG  721
Qy      657  -----  656
Db      722  AFGTVYKGLWPEGEKYKIPVAIKELREATSPANKELIDEAYVMASVDPNHYCRLLGIC  781
Qy      657  -----  656
Db      782  LTFVQVLITQLMPEGLLDYVREHKDNGISQYLLNMCVQIAGMYLIEDRLVHHDLAAR  841
Qy      657  -----DLG-----  659
Db      842  NLVKTPQHVKPTDFGLAKLLGAEEKYHAEGGKVPDKWMALESILHRTYHQSDWMSYG  901
Qy      660  -----PA-----  661
Db      902  VTVMELTFSKPYDGIIPASEISSILEKGERLPPRIGTIDVYMIWVKCMIDADSRRKF  961
Qy      662  -----SPLDSTFYRSLDEDDGDVLDAEEYLPQO  692
Db      962  RELIIEFSKMARDPQRLVYQGBERHMLBPPTDSNRYRBMDEEDMDVDVDAEDYLFIQO  1021
Qy      693  GFECPPDAPGAGGVYHHRHRSSTRSQGGDLTLGLEPSEBEAPRSFLAPSBGASGVDFDG  752
Db      1022  GFF-----SSPSTRKPTLLSLTSATSN--NS  1045
Qy      753  DLGGAAGKGLQSLPTDHPSPLORYSEDPVYLPSET--DGYVAPLTCSPQPEYVNOQDVR  810
Db      1046  TVACIDNGNGLQSCIKRSDSLQRTSSDPICALTEDSIDTFL-----VPEXYING-----  1099
Qy      811  POPPSREGPLPAARPGATL-----ERPKTSPKNGVYKDVAFAGAVENPEYL-TPQ  864
Db      1096  -----SWP-KRPASQVONPVYHNOPLNAPRSBDPHXD--PHSTAVGMPDELNTWQ  1144
Qy      865  GGAPQPHPPAPSPAPADNLVYMDQ-----DP-----PEGAPSPSTFGKTPPAE  908
Db      1144  -----PTCVNSTDPSPAHMAQSGSHQISLDNPDYQDDPEPKAEPKNGIFKGS-TAE  1195
Qy      909  NPEYL 913
Db      1194  NAEYL 1198

```

Search completed: January 13, 2003, 14:49:51  
Job time : 28.7768 secs



...

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run On: January 13, 2003, 14:44:55 ; Search time 17.4617 Seconds  
(without alignments)  
3919.881 Million cell updates/sec

Title: US-09-854-356-7

Perfect score: 3954

Sequence: 1 MELDALCRMGLLALLPRGA.....GFCCPDPAFGAGVHHRRH 712

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: PIR-73:\*  
2: PIR1:\*  
3: PIR2:\*  
4: PIR3:\*  
5: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3776	95.5	1255	1 A24571	protein-tyrosine k
2	3205	81.1	1260	1 TVRTMD	protein-tyrosine k
3	3185	80.6	1254	2 I48161	P-185 precursor -
4	1536	38.8	644	2 A36325	epidermal growth f
5	1534	38.8	1210	1 GQHUE	epidermal growth f
6	1533.5	38.8	1210	2 A53183	epidermal growth f
7	1509.5	38.2	1223	1 TVCHIV	epidermal growth f
8	1451	36.7	1308	2 A36223	epidermal growth f
9	1435	36.3	1342	2 A36223	kinase-related tra
10	1351.5	34.2	1339	2 J04387	epidermal growth f
11	1313.5	33.2	1166	1 S06142	protein-tyrosine k
12	1155	29.2	527	2 A42032	epidermal growth f
13	1014.5	25.7	843	2 A27131	epidermal growth f
14	718.5	18.2	1323	2 E88237	protein let-23 [im
15	718.5	18.2	1374	2 S70712	protein-tyrosine k
16	691.5	17.5	1330	1 GOFPE	epidermal growth f
17	669.5	16.9	1369	2 S70713	protein-tyrosine k
18	621	15.7	1717	1 A45558	epidermal growth f
19	384	9.7	1363	2 T43220	insulin-like growt
20	370	9.4	1300	2 A36502	insulin receptor-r
21	370	9.4	1382	1 INHUR	insulin receptor p
22	363.5	9.2	366	2 D45558	epidermal growth f
23	362.5	9.2	1383	2 A36057	insulin receptor p
24	361	9.1	1372	2 A34157	insulin receptor p
25	353	8.9	1477	2 T18534	protein-tyrosine k
26	348.5	8.8	540	2 B47417	insulin receptor-r
27	338	8.5	1268	2 B36502	insulin receptor-r
28	331	8.4	333	2 B45558	epidermal growth f
29	331	8.4	342	2 C45558	epidermal growth f

30	326	8.2	1390	2 T30346	insulin receptor -
31	324	8.2	1607	2 T43212	insulin-like growt
32	316	8.0	1371	2 A33837	insulin-like growt
33	311.5	7.9	1367	1 IGHUR1	insulin-like growt
34	307	7.8	2101	2 S57245	insulin receptor (
35	307	7.8	2148	1 A56081	insulin receptor -
36	267.5	6.8	1846	2 T42047	insulin receptor h
37	244.5	6.2	1548	2 S34583	serine proteinase
38	240	6.1	1299	2 T43251	furin (EC 3.4.21.7
39	235	5.9	1829	2 A48805	insulin-like growt
40	226.5	5.7	133	2 JH0803	tyrosine kinase re
41	208	5.3	1680	2 A43434	furin (EC 3.4.21.7
42	195.5	4.9	937	2 I53282	gene PACB4 protein
43	195	4.9	932	2 I52527	PACB4A - mouse (fr
44	194.5	4.9	1574	2 T13954	MEGF6 protein - ra
45	191.5	4.8	915	1 A48225	subtilisin-like pr

#### ALIGNMENTS

RESULT 1  
A24571  
protein-tyrosine kinase (EC 2.7.1.112) erbB2 precursor - human  
N:Alternate names: c-erb-B-2 protein precursor; kinase-related transforming protein e  
C:Species: Homo sapiens (man)  
C:Date: 25-Oct-1987 #sequence\_revision 06-Dec-1996 #text\_change 11-Jun-1999  
C:Accession: A24571; A25491; A44188; B44188; I59509; I57622  
R:Yamamoto, T.; Ikawa, S.; Akiyama, T.; Semba, K.; Nomura, N.; Miyajima, N.; Salto, T  
Nature 319, 230-234, 1986  
A:Title: Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growt  
A:Reference number: A24571; MUID:86118663; PMID:3003577  
A:Accession: A24571  
A:Molecule type: mRNA  
A:Residues: 1-1255 <YAM>  
A:Cross-references: GB:X03363; NID:g31197; PIDN:CAA27060.1; PID:g31198  
R:Semba, K.; Kamata, N.; Toyoshima, K.; Yamamoto, T.  
Proc. Natl. Acad. Sci. U.S.A. 82, 6497-6501, 1985  
A:Title: A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-erbB-1/epid  
A:Reference number: A25491; MUID:86016729; PMID:2995967  
A:Accession: A25491  
A:Molecule type: DNA  
A:Residues: 737-1031 <SEM>  
A:Cross-references: GB:M11767; NID:g182163; PIDN:AAA3808.1; PID:g553282  
R:Coussens, L.; Yang-Feng, T.L.; Liao, Y.C.; Chen, E.; Gray, A.; McGrath, J.; Seeburg  
Science 230, 1132-1139, 1985  
A:Title: Tyrosine kinase receptor with extensive homology to EGF receptor shares chro  
A:Reference number: A44188; MUID:86070181; PMID:2999974  
A:Accession: A44188  
A:Molecule type: DNA  
A:Residues: 740-910 <COU1>  
A:Cross-references: GB:M12036; NID:g183988; PIDN:AAA35978.1; PID:g183989  
A:Accession: B44188  
A:Molecule type: mRNA  
A:Residues: 1-517, 'RAL', 522, 'S', 524-654, 'V', 656-1169, 'A', 1171-1255 <COU2>  
A:Cross-references: GB:M11750; NID:g183986  
R:King, C.R.; Kraus, M.H.; Aaronson, S.A.  
Science 229, 974-976, 1985  
A:Title: Amplification of a novel v-erbB-related gene in a human mammary carcinoma.  
A:Reference number: I59509; MUID:85272597; PMID:2992089  
A:Accession: I59509  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 832-909 <REX>  
A:Cross-references: GB:L29395; NID:g459807; PIDN:AAA3809.1; PID:g459808  
R:Tel, M.; King, C.R.; Kraus, M.H.; Ulrich, A.; Schlesinger, J.; Glyol, D.  
Mol. Cell. Biol. 7, 2597-2601, 1987  
A:Title: Human HER2 (neu) promoter: evidence for multiple mechanisms for transcriptio  
A:Reference number: I57622; MUID:87286898; PMID:3039351  
A:Accession: I57622  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-191 <RAL>

A:Cross-references: GB:M16792; NID:9183983; PIDN:AAA58637.1; PID:9553332  
 C:Comment: Amplification and overexpression of this erbB-related gene occurs in about 30  
 C:Genetics:  
 A:Gene: GDB:ERBB2; NGL: NEU; HER-2  
 A:Cross-references: GDB:120613; OMIM:164870  
 A:Map position: 17q21.1-17q21.1  
 A:Introns: 25/1; 75/3; 147/1; 883/3  
 A>Note: the list of introns is incomplete  
 C:Function:  
 A:Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP  
 C:Superfamily: epidermal growth factor receptor; protein kinase homology  
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho  
 inase  
 F:1-21/Domain: signal sequence #status predicted <SIG>  
 F:22-1255/Product: protein-tyrosine kinase erbB2 #status predicted <MAT>  
 F:22-653/Domain: extracellular #status predicted <EXT>  
 F:70-304/Domain: EGF receptor extracellular domain repeat <EE1>  
 F:395-605/Domain: EGF receptor extracellular domain repeat <EE2>  
 F:654-675/Domain: transmembrane #status predicted <TM>  
 F:676-1255/Domain: intracellular #status predicted <INT>  
 F:718-983/Domain: protein kinase homology <KIN>  
 F:726-734/Region: protein kinase ATP-binding motif  
 F:68,124,187,259,530,571,629/Binding site: carbohydrate (Asn) (covalent) #status predict  
 F:686/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted  
 F:753/Active site: Lys #status predicted  
 F:1139,1221,1222,1248/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation)

Query Match 95.5%; Score 3776; DB 1; Length 1255;  
 Best Local Similarity 67.9%; Pred. No. 3.8e-244;  
 Matches 712; Conservative 0; Mismatches 0; Indels 336; Gaps 1;

```

OY 1 MELAALCRWGLLALLPPGAASSTOVCTGDMKRLPASPEHMLDMLHLYOGGVVGNL 60
Db 1 MELAALCRWGLLALLPPGAASSTOVCTGDMKRLPASPEHMLDMLHLYOGGVVGNL 60
OY 61 ELTYLPNNAISLFLQDIQEVGYVLAHNOYRQVPLQRLKIVKSTOLFENNALAVDNG 120
Db 61 ELTYLPNNAISLFLQDIQEVGYVLAHNOYRQVPLQRLKIVKSTOLFENNALAVDNG 120
OY 121 DPLNNTPTVPGASPGGLREQLSLFTLILKGVLIQNPOLCYODTILMDIFKKNQLA 180
Db 121 DPLNNTPTVPGASPGGLREQLSLFTLILKGVLIQNPOLCYODTILMDIFKKNQLA 180
OY 121 DPLNNTPTVPGASPGGLREQLSLFTLILKGVLIQNPOLCYODTILMDIFKKNQLA 180
Db 121 DPLNNTPTVPGASPGGLREQLSLFTLILKGVLIQNPOLCYODTILMDIFKKNQLA 180
OY 181 LTLIDTNRSRACHPCSPCKSGKSGSSSDCSLTFTVAGGACRCKGGLPTDCCHEQC 240
Db 181 LTLIDTNRSRACHPCSPCKSGKSGSSSDCSLTFTVAGGACRCKGGLPTDCCHEQC 240
OY 181 LTLIDTNRSRACHPCSPCKSGKSGSSSDCSLTFTVAGGACRCKGGLPTDCCHEQC 240
Db 181 LTLIDTNRSRACHPCSPCKSGKSGSSSDCSLTFTVAGGACRCKGGLPTDCCHEQC 240
OY 241 AACCTGKHSDDLACLFHNSGICELHCPALVYNTDTFESMPNDECRYTFGASCYACP 300
Db 241 AACCTGKHSDDLACLFHNSGICELHCPALVYNTDTFESMPNDECRYTFGASCYACP 300
OY 301 YNTLSTVSGCTVCPPLHNOEVAEDGTORCEKSKPCARCVGLGMEHLREVAIVSAN 360
Db 301 YNTLSTVSGCTVCPPLHNOEVAEDGTORCEKSKPCARCVGLGMEHLREVAIVSAN 360
OY 301 YNTLSTVSGCTVCPPLHNOEVAEDGTORCEKSKPCARCVGLGMEHLREVAIVSAN 360
Db 301 YNTLSTVSGCTVCPPLHNOEVAEDGTORCEKSKPCARCVGLGMEHLREVAIVSAN 360
OY 361 IOEFAGCKITFGSLAFPESEFGDPASNTAPLOPEQLQVETLEITGYLYISAMPDLP 420
Db 361 IOEFAGCKITFGSLAFPESEFGDPASNTAPLOPEQLQVETLEITGYLYISAMPDLP 420
OY 421 DLSYFQMLQYIRGILNNGAYSLTLOGLISWLGRLSRLREGSLALIHNTHTLCFVHTV 480
Db 421 DLSYFQMLQYIRGILNNGAYSLTLOGLISWLGRLSRLREGSLALIHNTHTLCFVHTV 480
OY 481 PMQOLFENPHOALLHTANRPEDECVGEGILACHOICARGHCHGPPPTCCVNCISOFLRQDEC 540
Db 481 PMQOLFENPHOALLHTANRPEDECVGEGILACHOICARGHCHGPPPTCCVNCISOFLRQDEC 540
OY 541 VEECRVQLQGLPREYVNAHCLPCHPECOPOGNGSYTCFPEADQCVAAHYKDPFCVARC 600
Db 541 VEECRVQLQGLPREYVNAHCLPCHPECOPOGNGSYTCFPEADQCVAAHYKDPFCVARC 600
OY 601 PSGVKPDLSTYPMWKPPDEEGACQPCPINCTHSCVDLDDKGCAPAEORASPLTSS 653
Db 601 PSGVKPDLSTYPMWKPPDEEGACQPCPINCTHSCVDLDDKGCAPAEORASPLTSS 653

```

```

Db 601 PSGVKPDLSTYPMWKPPDEEGACQPCPINCTHSCVDLDDKGCAPAEORASPLTSS 660
OY 654 ----- 653
Db 661 ILLVVLGVVFGILLIKRROOKIKRYTMRLLQETELVEPLTPSGAMPNOQRILKETEL 720
OY 654 ----- 653
Db 721 RKVKVLGSGAFGVYKGIWIPQGENVKIPALIKLRENTSPKANKETLDEAYVMAGVSP 780
OY 654 ----- 653
Db 781 YVSRLIGICLSTVQLVTLQIMPYCULLDHVENRGRSGQDLLMCMQIAGKMSYLEDVR 840
OY 654 ----- 653
Db 841 LVHRLDAARNVLYKSPNHVKITDPGLARLLDIDETEHADGGKVPILKMALESILARRPT 900
OY 654 ----- 653
Db 901 HQSDVMSYGVTVWELMTFGAKPYDGIPIAREIPDLLEKGERLPQPICTIDVYIMVYKCM 960
OY 654 ----- 684
Db 961 IDSECRPRFRELYSEFSRMAADPQREVVIONEDLGASPLDSTFYSLLLEDMDGLVDA 1020
OY 685 EEYLVPOQGFPCPDPAAGAGVHHRHR 712
Db 1021 EEYLVPOQGFPCPDPAAGAGVHHRHR 1048

```

## RESULT 2

## TVRTNU

protein-tyrosine kinase (EC 2.7.1.112) neu precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #cext\_change 11-Jun-1999

C:Accession: A24562; A61204

R:Bargmann, C.I.; Hung, M.C.; Weinberg, R.A.

Nature 319, 226-230, 1986

A:Title: The neu oncogene encodes an epidermal growth factor receptor-related protein

A:Reference number: A24562; MUID:86118662; PMID:3945311

A:Accession: A24562

A:Molecule type: mRNA

A:Residues: 1-1260 <BAR>

A:Cross-references: EMBL:X03362; NID:956745; PIDN:CAA27059.1; PID:956746

R:Masu, T.; Mann, A.M.; Macatee, T.L.; Garland, E.M.; Okamura, T.; Smith, R.A.; Cohe

A:Title: Direct DNA sequencing of the rat neu oncogene transmembrane domain reveals n

2-thiazolyl]formamide or N-methyl-N-nitrosourea.

A:Reference number: A61204; MUID:92035293; PMID:1682063

A:Accession: A61204

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 637-663, 'V', 665-702 <MAS>

A>Note: authors translated the codon GCA for residue 25 as Val

C:Genetics:

A:Gene: neu

C:Superfamily: epidermal growth factor receptor; protein kinase homology

C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phos

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-1260/Product: protein-tyrosine kinase neu #status predicted <MAT>

F:658-680/Domain: transmembrane #status predicted <TM>

F:723-988/Domain: protein kinase homology <KIN>

F:731-739/Region: protein kinase ATP-binding motif

F:71,191,263,535,576,634/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:661/Binding site: phosphate (Thr) (covalent) #status predicted

F:758/Active site: Lys #status predicted

F:1139,1227,1253/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 81.1%; Score 3205; DB 1; Length 1260;  
 Best Local Similarity 57.9%; Pred. No. 5e-206;  
 Matches 608; Conservative 34; Mismatches 70; Indels 338; Gaps 3;

```

Oy 1 MELAALCRWGLLALLPFGAASVOYCTGDMKRLRASPETHLDMIRHLRYOGCQVVGNTL 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4 MELAAMCRWGLLALLPFIAGTOYCTGDMKRLRASPETHLDMIRHLRYOGCQVVGNTL 63
Oy 61 ELTYLPTNASSFLDIDIEVOGYVLIANNOVQVLPQRLRIYRGQLFEDNTALAVLNG 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 64 ELTYVPANASSFLDIDIEVOGYVLIANNOVQVLPQRLRIYRGQLFEDNTALAVLNG 123
Oy 121 DPLNNTPTVT-GASPGGRLRELQRLSLTELKGVLLIQNRPOCYODTILMKDIFHKNNQL 179
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 124 DPQDNVAASPTGRTPEGRLEQLRLSLTELKGVLLIRGNPOLCYODMVLKVFRRNNQL 183
Oy 180 ALTLIDTNRSRACHPCSPMKSGRCWGESSEDCOSLTRVNCAGCARKGRLPTDCCHQ 239
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 184 APVDIDTNRSRACHPCAPARCKDNHCWGESPECCQLITGICTSGCARKGRLPTDCCHQ 243
Oy 240 CAAGCTGPKHSDCLACLFHNHSGICELHCPALVTYNTDFTESMPNREGRTFGASCVTAC 299
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 244 CAAGCTGPKHSDCLACLFHNHSGICELHCPALVTYNTDFTESMHPNREGRTFGASCVTAC 303
Oy 300 PYNVLTSTVSGCTLVCPRLNNOVTAEDGTQRCCKSKPCARVCYGLGMHLREVRATVSA 359
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 304 PYNVLTSTVSGCTLVCPRLNNOVTAEDGTQRCCKSKPCARVCYGLGMHLREVRATVSD 363
Oy 360 NIOEPAGCKIIFGSLAFIPESFDGPASNTAPLOEQLQVFTLEITGVLISAMPDL 419
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 364 NIOEPAGCKIIFGSLAFIPESFDGPASNTAPLOEQLQVFTLEITGVLISAMPDL 423
Oy 420 PDLVFQNLQVIRGRILHNGAVSLTLQIGISMLGLRSLRELGSGLALIHNTHTLCFYHT 479
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 424 RDLVFQNLQVIRGRILHNGAVSLTLQIGISMLGLRSLRELGSGLALIHNTHTLCFYHT 483
Oy 480 VPMQDLFRNPQALHTANRPEDF-CVGEGLACHOLCARHCWGPCPQCVNCSQFLKQ 538
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 484 VPMQDLFRNPQALHTANRPEDF-CVGEGLACHOLCARHCWGPCPQCVNCSQFLKQ 543
Oy 539 ECVEECRVLOGLPREVYVARNHCLPCHPECOPONGSVTCGPRADOCVCAHKKDPFCA 598
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 544 ECVEECRVLOGLPREVYVARNHCLPCHPECOPONGSVTCGPRADOCVCAHKKDPFCA 603
Oy 599 RCPGVKBDLSYMPIMKPEDEGACOPCPINCTHSCVDLDDKGCFAEDRASPLT 652
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 604 RCPGVKBDLSYMPIMKPEDEGACOPCPINCTHSCVDLDDKGCFAEDRASPLT 663
Oy 653 ----- 652
Db 664 EGVLLFLVVVVGILIKRRROKIRKTYMRLLQETELVEPLTPSGAMPNQAOMRIKET 723
Oy 653 ----- 652
Db 724 ELRKVKVLGSGAFVYVGIWIPDGENVKIPVAIKVLEBNTSPRANKETLDEAVYMAVG 783
Oy 653 ----- 652
Db 784 SPYVSRILGICLTSTVQLTQMLPYGCLLDHVRHGRRLSGODLLMWCQIAKMSYLE 843
Oy 653 ----- 652
Db 844 VRLVHRDLAARNVLKSPNHVKITDGLARLLIDETEVHADGKVPITKMALESILRRR 903
Oy 653 ----- 652
Db 904 FTHQSDVMSYGVTVWELMTFGAKPYDGIAPAREIPDLKEGELRPPICITIDVYIMKXC 963
Oy 653 -----SONEDLGPASPLDSTFYRSLLEDDMDGLV 682
Db 964 WMIDSECRPRRELVSSEFSRMAADPQRFVYVIONEDLGSSPMDSTFYRSLLEDDMDGLV 1023
Oy 683 DAEETLVPOGCFPCPDPAAGAGMHHNR 712
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1024 DAEETLVPOGCFPCPDPAAGAGMHHNR 1053

```

```

RESULT 3
p-185 precursor - golden hamster
148161
C:Species: Mesocricetus auratus (golden hamster)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Jun-1999
C:Accession: 148161
R:Nakamura, T.; Ushijima, T.; Ishizaka, Y.; Nago, M.; Arai, M.; Yamazaki, Y.; Ishika
Gene 140, 251-255, 1994
A>Title: Cloning and activation of the Syrian hamster neu proto-oncogene.
A:Reference number: 148161; MUID:94193007; PMID:7908275
A:Accession: 148161
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1254 <RSS>
A:Cross-references: GB:D16295; NID:g493236; PIDN:BA03801.1; PID:g747595
C:Genetics:
A:Gene: neu
C:Superfamily: epidermal growth factor receptor; protein kinase homology
C:Keywords: Atp
F:718-983/Domain: protein kinase homology <KIN>
F:726-734/Region: protein kinase Atp-binding motif

Query Match 80.6%; Score 3185; DB 2; Length 1254;
Best Local Similarity 57.5%; Pred. No. 1.1e-204;
Matches 603; Conservative 36; Mismatches 73; Indels 336; Gaps 1.

Oy 1 MELAALCRWGLLALLPFGAASVOYCTGDMKRLRASPETHLDMIRHLRYOGCQVVGNTL 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MELAAMCRWGLLALLPFIAGTOYCTGDMKRLRASPETHLDMIRHLRYOGCQVVGNTL 60
Oy 61 ELTYLPTNASSFLDIDIEVOGYVLIANNOVQVLPQRLRIYRGQLFEDNTALAVLNG 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 ELTYVPANASSFLDIDIEVOGYVLIANNOVQVLPQRLRIYRGQLFEDNTALAVLNG 120
Oy 121 DPLNNTPTVTGASPGGRLRELQRLSLTELKGVLLIQNRPOCYODTILMKDIFHKNNQLA 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 DPLNNTPTVTGASPGGRLRELQRLSLTELKGVLLIRGNPOLCYODTILMKDIFHKNNQLA 180
Oy 61 ELTYLPTNASSFLDIDIEVOGYVLIANNOVQVLPQRLRIYRGQLFEDNTALAVLNG 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 ELTYVPANASSFLDIDIEVOGYVLIANNOVQVLPQRLRIYRGQLFEDNTALAVLNG 120
Oy 121 DPLNNTPTVTGASPGGRLRELQRLSLTELKGVLLIQNRPOCYODTILMKDIFHKNNQLA 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 DPLNNTPTVTGASPGGRLRELQRLSLTELKGVLLIRGNPOLCYODTILMKDIFHKNNQLA 180
Oy 181 LTLIDTNRSRACHPCSPMKSGRCWGESSEDCOSLTRVNCAGCARKGRLPTDCCHQ 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 LTLIDTNRSRACHPCSPMKSGRCWGESSEDCOSLTRVNCAGCARKGRLPTDCCHQ 240
Oy 181 PVDIDTNRSRACHPCAPARCKDNHCWGESPECCQLITGICTSGCARKGRLPTDCCHQ 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 PVDIDTNRSRACHPCAPARCKDNHCWGESPECCQLITGICTSGCARKGRLPTDCCHQ 240
Oy 241 AAGCTGPKHSDCLACLFHNHSGICELHCPALVTYNTDFTESMPNREGRTFGASCVTAC 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 AAGCTGPKHSDCLACLFHNHSGICELHCPALVTYNTDFTESMPNREGRTFGASCVTAC 300
Oy 301 YNVLTSTVSGCTLVCPRLNNOVTAEDGTQRCCKSKPCARVCYGLGMHLREVRATVSA 360
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 YNVLTSTVSGCTLVCPRLNNOVTAEDGTQRCCKSKPCARVCYGLGMHLREVRATVSA 360
Oy 361 IOEFAGCKKIFGSLAFIPESFDGPASNTAPLOEQLQVFTLEITGVLISAMPDL 420
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 IOEFAGCKKIFGSLAFIPESFDGPASNTAPLOEQLQVFTLEITGVLISAMPDL 420
Oy 421 DLSVFQNLQVIRGRILHNGAVSLTLQIGISMLGLRSLRELGSGLALIHNTHTLCFYHT 480
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 DLSVFQNLQVIRGRILHNGAVSLTLQIGISMLGLRSLRELGSGLALIHNTHTLCFYHT 480
Oy 481 PWDOLFNRNPQALHTANRPEDF-CVGEGLACHOLCARHCWGPCPQCVNCSQFLKQ 538
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 481 PWDOLFNRNPQALHTANRPEDF-CVGEGLACHOLCARHCWGPCPQCVNCSQFLKQ 538
Oy 541 VEECRVLOGLPREVYVARNHCLPCHPECOPONGSVTCGPRADOCVCAHKKDPFCA 600
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 541 VEECRVLOGLPREVYVARNHCLPCHPECOPONGSVTCGPRADOCVCAHKKDPFCA 600
Oy 601 PSQVAPDLSYMPIMKPEDEGACOPCPINCTHSCVDLDDKGCFAEDRASPLT 653
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 601 PSQVAPDLSYMPIMKPEDEGACOPCPINCTHSCVDLDDKGCFAEDRASPLT 653
Oy 654 ----- 653
Db 661 ILFLVIGVVVGIILIKRRROKIRKTYMRLLQETELVEPLTPSGAMPNQAOMRIKETEL 720

```

```

QY 654 ----- 653
Db 721 RKVVLGSGAFGYKGIWIPDGENVKIPVAKVLRRENTSPKANKELLDEAYVWAGLSP 780
QY 654 ----- 653
Db 781 YSRHLGICLITVQVLTOLMPYGLLDHVEHRRGLGSODLLMWQVQIAKMSYLEDV 840
QY 654 ----- 653
Db 841 LVHMDLARNVLYKSPNHVKTITDFGLARLLDIDETYNADGKVPDKMALESLIRRET 900
QY 654 ----- 653
Db 901 HQSDWXYGYTVWELMTFGAKPYDIPARELPDLEKGERLPQPPICIDVYIMVNCWM 960
QY 654 ----- 684
Db 961 IDSECRPRFRELVSFSEMRMARDPQRFVYVIONEDLGSSPLDSTFYSRLLEDMDGLDVA 1020
QY 685 EEYLVPOGFEPCPDAPGAGMHHRR 712
Db 1021 EEYLVPOGFEPCPDAPGAGMHHRR 1048

RESULT 4
A:36325
epidermal growth factor receptor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 25-Jan-1991 #sequence, revision 25-Jan-1991 #text, change 10-Oct-1997
C:Accession: A36325
R:Petich, L.A.; Harris, J.; Raymond, V.W.; Blasband, A.; Lee, D.C.; Earp, H.S.
Mol. Cell. Biol. 10, 2973-2982, 1990
A:Title: A truncated, secreted form of the epidermal growth factor receptor is encoded b
A:Reference number: A36325; MUID:90258886; PMID:2342466
A:Accession: A36325
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-644 <P>
A:Cross-references: GB:M37394
C:Superfamily: epidermal growth factor receptor; protein kinase homology
C:Keywords: alternative splicing; ATP; growth factor receptor

Query Match 38.8%; Score 1536; DB 2; Length 644;
Best Local Similarity 44.6%; Pred. No. 6e-95;
Matches 287; Conservative 96; Mismatches 234; Indels 24; Gaps 8;

QY 3 LAALCRWGLLALLPGA-ASTQVCTGTDMKRLPASPETHLDMRLHYOCQOVVGNLE 61
Db 15 LAALCAAG-----GALKEKVKCGTNSRLTQLGTFEDHFLSLQRMFNCEVYVNL 66
QY 62 LTYLPTNASTFLDIDDEVGYVLIANNQVAPLQRIKRYGQLEPEDNATLAVLDNGD 121
Db 67 ITTVQRYVDLSFLKTIDEVAGYVLIATNTERIPLENIQIIRGNALYENTYALVLSN-- 124
QY 122 PLNNTTPTVGTASPGRELOLRSLEIKGGVLIQORNPOLCYODPIIMKODIFHNODAL 181
Db 125 -----YGNKTGLKELPRKIQEILIGAVRESNPNILCNMETIQRDLY-ODVELSN 175
QY 182 TLIDTNRN-RACHPCSPMKSGKRCMGSESDQSLTRTVCAAGCA-RCKGFLPTDCCHEQ 239
Db 176 MSMDVQHLNGCPCDPSRCMGSGMBEHCQKLTIKACQOGRSGRSGSPDCCINQ 235
QY 240 CAAGCTPRKISDCLACHFNHSGICEHLCPALVYNTDFESMPNPGRYTGGASVYAT 299
Db 236 CAAGCTPRPSDCLVCHRFREDEATCKDTCPLMLYNPTTYQMDVNPBGKYSFGATCYVKK 295
QY 300 PYNTLSIDVSGCTLVCLHNGEYTAEDGTQRCCKSPCAVYCGICGMEHLREVRATSA 359
Db 296 PRNVTVVDHSCVRACPDYIEV-EEQGVSSCKKCDKPCRRVNGVIGERKDLISNAT 354
QY 360 NIOFACCKIKFGSLATPESFDGDPASNTAPLQPEQVQVETLEITGYLYISAMPDSL 419

```

```

Db 355 NIKHFKYCTAISGLDHLILPVAFKGDSEFTRPPLDPRELEILTKVKEITGFLLIQAMPENW 414
QY 420 PDLSEFQNLQYIRGRILHNGAYSLTLOGLSWLGSLNSRLSGALLIHHNHLCEVHT 479
Db 415 TDLAFENLELIRGTRQHGQFSLAVVGLNTSLGSLSLKEISGDDVYISGNRLCIANT 474
QY 480 VPMQLFRNPQALHTANRPEDCEVGEGLACHQLCARGHCWGPGTQVNCQSOFLGQE 539
Db 475 INMKKLFETPRQKTKINNRAEKCKAFNHYCNFLGSESGCMGEPIDCVSCQVNSGR 534
QY 540 CVBECRYLQGLPREYVNAHCLPCHPECOQNGSVTCGFPADQCVACAHYKDPFCVAR 599
Db 535 CVDKCNLIEGPRFREVENSECICQHPCLPOTMNTITGREGPDNCRICAHYVDGPHCVKT 594
QY 600 CPSCGVKPDLSYMPIMKPFDEGACQPPCINCTHSCVDLDKGC 642
Db 595 CPSCIMGBNNTL-VWKFADANNVCHLANCTYGCAGPGLKGC 636

RESULT 5
GQHUH
epidermal growth factor receptor precursor - human
N:Contains: protein-tyrosine Kinase (EC 2.7.1.112) erbB
C:Species: Homo sapiens (man)
C>Date: 15-Nov-1984 #sequence, revision 27-Nov-1985 #text, change 11-Jun-1999
C:Accession: A00641; A25772; S30024; A38672; A43615; A23062; A05281; A60143;
R:Ullrich, A.; Coussens, L.; Hayflick, J.S.; Dull, T.J.; Gray, A.; Tam, A.W.; Lee, J.
Tg, P.H.
Nature 309, 418-425, 1984
A:Title: Human epidermal growth factor receptor cDNA sequence and aberrant expression
A:Reference number: A00641; MUID:84219729; PMID:6328312
A:Accession: A00641
A:Molecule type: mRNA
A:Residues: 1-1210 <L>
A:Cross-references: EMBL:X00588; NID:q31113; PIDN:CAA25240.1; PID:q757924
A:Note: the authors translated the codon AAG for residue 540 as Asn
R:Ishii, S.; Xu, Y.; Stratton, R.H.; Roe, B.A.; Merlino, G.T.; Pastan, I.
Proc. Natl. Acad. Sci. U.S.A. 82, 4920-4924, 1985
A:Title: Characterization and sequence of the promoter region of the human epidermal
A:Reference number: A25772; MUID:85270436; PMID:2991899
A:Accession: A25772
A:Molecule type: DNA
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-29 <H>
A:Cross-references: GB:M11234; NID:q181981; PIDN:AAA2370.1; PID:q553272
R:Hailey, J.; Whitte, N.; Bennett, P.; Kitchington, D.; Ullrich, A.; Waterfield, M.
Oncogene Res. 1, 375-396, 1987
A:Title: The human EGF receptor gene: structure of the 110 kb locus and identificatio
A:Reference number: S30024; MUID:88217333; PMID:3329716
A:Accession: S30024
A:Molecule type: DNA
A:Residues: 1-29 <HA>
A:Cross-references: EMBL:X06370; NID:q31118; PIDN:CAA29668.1; PID:q31119
R:Hailey, J.D.; Waterfield, M.D.
J. Biol. Chem. 266, 1746-1753, 1991
A:Title: Contributory effects of de Novo transcription and premature transcript termi
A:Reference number: A38672; MUID:91107677; PMID:1988448
A:Accession: A38672
A:Molecule type: DNA
A:Residues: 1-29 <H>
A:Cross-references: GB:M38425; NID:q181977; PIDN:AAA63171.1; PID:q553271
A:Experimental source: carcinoma cell line A431-7
R:Xu, Y.; Ishii, S.; Clark, A.J.L.; Sullivan, M.; Wilson, R.K.; Ma, D.P.; Roe, B.A.;
Nature 309, 806-810, 1984
A:Title: Human epidermal growth factor receptor cDNA is homologous to a variety of RN
A:Reference number: A00642; MUID:84245835; PMID:6330563
A:Accession: A00642
A:Molecule type: mRNA
A:Residues: 'RCAMRRA', 150-187, 'KSYTQAV', 195, 'W', 197, 'A', 199-222, 'S', 224-304, 'RA', 307-
', 798-799, 'TD', 802-811, 'R', 813-942 <X>
A:Experimental source: A431 human carcinoma cells, which have large numbers of EGF re
R:Lin, C.R.; Chen, W.S.; Kuilger, W.; Stolarsky, L.S.; Weber, W.; Evans, R.M.; Verna,

```



Science 224, 843-848, 1984  
A:Title: Expression cloning of human EGF receptor complementary DNA: gene amplification  
A:Reference number: A43615; MUID:84196372; PMID:6326261  
A:Accession: A43615  
A:Molecule type: mRNA  
A:Residues: 713-964 <LIN>  
A:Experimental source: epidermoid carcinoma cell line A431  
R:Slamen, F.A.; Gope, M.L.; Schulz, T.2.; Wright, D.A.; Carpenter, G.; O'Malley, B.W.  
Biochem. Biophys. Res. Commun. 124, 125-132, 1984  
A:Reference number: A23062; MUID:85046483; PMID:6093780  
A:Accession: A23062  
A:Molecule type: mRNA  
A:Residues: 1028-1210 <SIM>  
R:Weber, M.; Gill, G.N.; Speiss, J.  
Science 224, 294-297, 1984  
A:Reference number: A05281; MUID:84112183; PMID:6324343  
A:Accession: A05281  
A:Molecule type: protein  
A:Residues: 25-30, 'S', 32-51, 454-467 <WEB>  
R:Russol, M.W.; Lukas, T.J.; Cohen, S.; Staros, J.V.  
J. Biol. Chem. 260, 5205-5208, 1985  
A:Title: Identification of residues in the nucleotide binding site of the epidermal growth factor receptor  
A:Reference number: A60143; MUID:85182650; PMID:2985580  
A:Accession: A60143  
A:Molecule type: protein  
A:Residues: 740-744, 'X', 746-747 <RUS>  
R:Motczkowski, B.; Mosig, G.; Cohen, S.  
Nature 309, 270-273, 1984  
A:Title: ATP-stimulated interaction between epidermal growth factor receptor and superoxide dismutase  
A:Reference number: A38023; MUID:84191554; PMID:6325948  
A:Accession: A38023  
A:Contents: annotation; receptor activity  
A:Note: the EGF receptor (and other tyrosine kinases) can nick double-stranded DNA  
R:Chen, W.S.; Lazar, C.S.; Lund, K.A.; Welsh, J.B.; Chang, C.P.; Walton, G.M.; Der, C.J.  
Cell 59, 33-43, 1989  
A:Title: Functional independence of the epidermal growth factor receptor from a domain that mediates internalization  
A:Reference number: A33331; MUID:90003233; PMID:2790960  
A:Accession: A33331  
A:Contents: annotation; internalization signal  
C:Comment: Binding of EGF to the receptor leads to internalization of the EGF-receptor complex  
C:Genetics:  
A:Gene: GDB:EGFR  
A:Cross-references: GDB:120610; OMIM:131550  
A:Map position: 7p12.3-7p12.1  
A:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phosphotyrosine; signal sequence; status predicted <SIG>  
C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phosphotyrosine; signal sequence; status predicted <SIG>  
F:1-24/Domain: signal sequence #status predicted <SIG>  
F:25-1210/Product: EGF receptor #status predicted <EXT>  
F:25-645/Domain: extracellular domain repeat <EXT>  
F:75-300/Domain: EGF receptor extracellular domain repeat <EXT>  
F:390-600/Domain: EGF receptor extracellular domain repeat <EXT>  
F:668-668/Domain: transmembrane #status predicted <TM>  
F:668-1210/Domain: intracellular #status predicted <INT>  
F:710-915/Domain: protein kinase homology <KIN>  
F:718-726/Region: protein kinase ATP-binding motif  
F:999-1046/Region: coated-pit mediated internalization signal  
F:1047-1210/Region: inhibitory  
F:128,175,352,413,444,528,603/binding site: carbohydrate (Asn) (covalent) #status predicted  
F:745/Active site: lys #status experimental

Query Match 38.88; Score 1534; DB 1; Length 1210;  
Best Local Similarity 42.98; Pred. No. 1,66-94;  
Matches 283; Conservative 109; Mismatches 240; Indels 28; Gaps 6;

Qy 11 LLLALLPAGAA--STVOVGTGDMKRLPASPETHIDMKRLHYOCQOVQVNLGLTYLPTN 68  
Db 14 LLAALCPSPRALLEKKVQCGTSNKLTQIGREDHLSLQRMFNCEVVLGLNLETITYVGRN 73  
Qy 69 ASLSFIDIDQEVQGVYLAHNOVQVPLQRLRIYVGTQLFEDFNALAVLNGDPLNNTTP 128  
Db 74 YDLSTFKIQVAGVLAATVTERIPLENQIRGNMYENSTALAVLSYD----- 126  
Qy 129 VTGASPGGLRELQRLSTREILKGGVLIQNRDOLCYDTIIMKIDIFHKNNQALATLITDNR 188  
Db 127 ---ANKTGILKELPMRNLOEILHGAIVRFSSNPALCNVESIQNRDIIVSDPLSNMSMDQNH 183

Qy 189 SRACHPCSPMKGSGKSGWSESSSDOSLRTVCAGGCA-RCKGPIPTDCHEQCAAGCTGP 247  
Db 194 LGSCKQCDPSCPNSGSCWAGEBNCOKLTKITICAOQCSGCRKRSDDCHQCAAGCTGP 243  
Qy 248 KHSDDLACLNHNSGICLHCPALVTYNTDFEASPNREGRTFGASCVTACPYNYLSTD 307  
Db 244 RESDCLVCRKRDEATCDTQCPPLMLYPTTYQMDVNEBGRKSRATCYKRCPRYVYTD 303  
Qy 308 VGSCTLVCPNLNQEYTAEDGTFORCEKSCPCARVCYGLGMEHLREVAVTSANTIOEFAQC 367  
Db 304 HGSCVACGADSYEM-EDDGVKCKCKEPCRCVNCNGIGIEFKDSLSINNTIKHFKNC 362  
Qy 368 KIRGSLAFLESPFDPDASNTAPLQRPOLQVETLEITGYLTSAMPDPLDLSYQN 427  
Db 363 TSIISGDLHLPLVAFRGDSFTHTPPLDPDLILTKVKKEITGFLIQAMPENRTDLHAFEN 422  
Qy 428 LOVIRGRLIHNGAVSLTQGLGISWLGRLSRELGGSLALHNNHLCFHTVPDOLF 487  
Db 423 LEITGRTRKQHGQPSLAVSLNTISLGRSLKEISDGVYISGNKNTCYANTIMKKLFG 482  
Qy 488 NPHQALLHTANRPDECEVGEGLACHQLCARHGWGPGPTQCVNCSQFLRGQVEECRYL 547  
Db 483 TSGQRTKLIISNKGNSCATGQVCHALCSPEGCMGPEPRDCVSCRNVRSGRECVDKKTL 542  
Qy 548 OGPRELYNARHCLPHEPCOPQNSVTCFGEADQCAACHYKDPFCVACRPGCVAPD 607  
Db 543 EGEPREFENSECICQHCCELPQAMNITCTGRGPNCICQACHYIDGRCVCTCPAGVAGE 602  
Qy 608 LSYPMYKFPDEGACQCPINCFHSCYVDLDKGCPCAPQARASPLTSQNEDGASPLDST 667  
Db 603 NMTL-VMKYADAGVCHLCHRNCTYGCYPLGEGCPN-----GPKIPSIAT 648

RESULT 6  
A53183  
A:Title: Epidermal growth factor receptor precursor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 06-Jan-1995 #sequence revision 06-Jan-1995 #text change 18-Jun-1999  
C:Accession: A53183; A43818; S24942; A28941; S45325; I49643  
R:Luetkeke, N.C.; Phillips, H.K.; Qiu, T.H.; Copeland, N.G.; Earp, H.S.; Jenkins, N.A.  
Genes Dev. 8, 399-413, 1994  
A:Title: The mouse waved-2 phenotype results from a point mutation in the EGF receptor  
A:Reference number: A53183; MUID:94110986; PMID:8125255  
A:Accession: A53183  
A:Molecule type: mRNA  
A:Residues: 1-1210 <LUE>  
A:Cross-references: GB:U03425  
R:Aviv, A.; Lax, I.; Ullrich, A.; Schlessinger, J.; Givol, D.; Morse, B.  
Oncogene 6, 673-676, 1991  
A:Title: Comparison of EGF receptor sequences as a guide to study the ligand binding  
A:Reference number: A43818; MUID:91232866; PMID:2030916  
A:Accession: A43818  
A:Molecule type: mRNA  
A:Residues: 1-714 <AVI>  
A:Cross-references: GB:X59698  
R:Eisinger, D.P.; Serrero, G.  
Submitted to the EMBL Data Library, June 1992  
A:Accession: S24942  
A:Molecule type: mRNA  
A:Residues: 969-971, 'K', 973-1115, 'D' <EIS>  
A:Cross-references: EMBL:Z12608  
R:Heisermann, G.U.; Gill, G.N.  
J. Biol. Chem. 263, 13152-13158, 1988  
A:Title: Epidermal growth factor receptor threonine and serine residues phosphorylate  
A:Reference number: A28941; MUID:88330814; PMID:3138233  
A:Accession: A28941  
A:Molecule type: protein  
A:Residues: 689-694, 'X', 696-704, 'V', 706-707, 989-992, 'XX', 995-996, 'X', 998-1000, 1002-10  
R:Rubb, M.L.; Dunn, A.R.; Alexander, W.S.  
Submitted to the EMBL Data Library, April 1994  
A:Description: The complete cDNA sequence of the Mouse Epidermal Growth Factor Recept

A:Reference number: S45325  
 A:Accession: S45325  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1971, 'K', 973-1210 <VER>  
 A:Cross-references: EMBL:X78987; NID:g488830; PIDN:CAA5587.1; PID:g488831  
 R:Paria, B.C.; Das, S.K.; Andrews, G.K.; Day, S.K.  
 Proc. Natl. Acad. Sci. U.S.A. 90, 55-59, 1993  
 A:Title: Expression of the epidermal growth factor receptor gene is regulated in mouse  
 A:Reference number: I49643; MIMD:93126380; PMID:7678348  
 A:Accession: I49643  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 12-20, 22-132 <RES>  
 A:Cross-references: GB:I06864; NID:g193001; PIDN:AA53029.1; PID:g567201  
 C:Genetics:  
 A:Gene: EGFR  
 C:Superfamily: epidermal growth factor receptor; protein kinase homology  
 C:Keywords: ATP; growth factor receptor; kinase-related transforming protein; phosphop  
 F:1-24/Domain: signal sequence #status predicted <SIG>  
 F:648-670/Domain: transmembrane #status predicted <TM>  
 F:712-977/Domain: protein kinase homology <KIN>  
 F:720-728/Region: protein kinase ATP-binding motif  
 F:680,695/Binding site: phosphate (Thr) (covalent) #status experimental  
 F:697,1070,1071/Binding site: phosphate (Ser) (covalent) #status experimental  
 F:993/Binding site: (or 997) phosphate (Ser) (covalent) #status experimental  
 F:1028/Binding site: (or 1030 or 1032) phosphate (Ser) (covalent) #status experimental  
 F:1197/Binding site: phosphate (Tyr) (covalent) #status experimental

Query Match 38.8%; Score 1533.5; DB 2; Length 1210;  
 Best Local Similarity 44.4%; Pred. No. 1.8e-94;  
 Matches 284; Conservative 99; Mismatches 233; Indels 23; Gaps 7;

QY 11 LLLALLPFGAA--STOYCTGTDMKRLRPASPEHLDMLRLHYGCCOYVGGLEITYPTN 68  
 Db 14 LTRALCAAGALEERKVCQCGSNRLTQLGFEDHFLSLQRYNNCEVVLGELITYVQRN 73  
 QY 69 ASLSFLDIOEVQGYVLIANQVROYQLRLRYRTQLEFEDNALVALVDNGDPLNNTTP 128  
 Db 74 YDSEFLKTIOEVAGYVLIANTVERIPLENQILRGNALENTYALAILSN----- 124  
 QY 129 VTGASPEGLRELQRLSTLEILKGVLIQRNPQLCYODTILMKDI----FHKNNOLATLI 184  
 Db 125 -YGNRTGLRELPMRNQETLIGAVRSPNNPILCNMDTIQRDIYQVFNMSMDL--- 180  
 QY 185 DTNRSRACHPCSPKCSRCMGESSEDCQSLRTVCAGCA-RCKGPLPDDCCHEGCAAG 243  
 Db 181 -QSHPSPCPKDPCSPNGSCMGEGGECQKLTITCAQOCSHRCGRSPSCCHNCAAG 239  
 QY 244 CTGKHSDCIACLFHNSGICELHCPALVTYNTDFESMPNPEGRYFGASCYACPYNY 303  
 Db 240 CTGRESDCIACVCFQDEARCKDTCPLMLYNTTYQMDVNPBEKYSFGATCYVKKCRNY 299  
 QY 304 LSTDVGSCTIVCPLIHQNEVTAEDGTQRCERKCSKPCARVCYGLGMEHLREYRAVTSANIOE 363  
 Db 300 VVVDHGCVACVAGDYEV--EEDGIRCKKCDGCRKVCNGSIGEFKDTLSINATIKH 358  
 QY 364 FAGCKKIFGSLAPLPESFDDPASNTAPLQPELOVFEETLEITGYLISAMPDLSLDLS 423  
 Db 359 FKCTALISGLHILPAVKDSFTRPRLDRELLEILTKYKELTGILLIADMPDNMVDLH 418  
 QY 424 VFONLOVIRIRLIHNGAYSTLLOGISLWGLRSLRELGLSGLAILHNTHLCFEYTPMD 483  
 Db 419 AFENLEIRRTQKHQGFSLAVVGLNTISGLRSLKEISDGVITISGRNICYANTINMK 478  
 QY 484 QLEFRNPHOALLHTANPEDECVGEGLACHQLCARGHCWGPGPTQCVNCSQFLRGQECVEE 543  
 Db 479 LKGTIPNOKTKIMNNRAEKDKAVNHVCNPLCSSEGCWGPEPRCVCQNNVSRGECVEK 538  
 QY 544 CRLVGLGIPREYVANAHLCPHREPCQPNQSVYTCGPPADOCVACAHKDPFCYARPSG 603  
 Db 539 CNLEGPREFVENSECIOQHPECLPQAMNITCTGRRGPDNCIOCAHYIDGPHCVKTCBAG 598

QY 604 VKPDLSTYPIWKPEDEGACPCPINCTHSCVDLDDKGC 642  
 Db 599 IMGENNRL-VWKYADANNVCHLANCTYGCAGPGLQGC 636

# RESULT 7

## IVCHLV

epidermal growth factor receptor precursor - chicken

N:Contins: protein-tyrosine kinase (EC 2.7.1.112) erbB

C:Species: Gallus gallus (chicken)

C>Date: 28-Feb-1986 #sequence\_revision 05-May-1995 #extl\_change 04-Feb-2000

C:Accession: A27720; A00643

R:Lax, I.; Johnson, A.; Hawk, R.; Sap, J.; Bellot, F.; Winkler, M.; Ullrich, A.; Venn

Mol. Cell. Biol. 8, 1970-1978, 1988

A:Title: Chicken epidermal growth factor (EGF) receptor: cDNA cloning, expression in

A:Reference number: A27720; MIMD:88261272; PMID:3260329

A:Accession: A27720

A:Molecule type: mRNA

A:Residues: 1-1223 <LAX>

A:Cross-references: GB:M20386

R:Nilsen, T.W.; Maroney, P.A.; Goodwin, R.G.; Rottman, F.M.; Crittenden, L.B.; Raines

Cell 41, 719-726, 1985

A:Title: c-erbB activation in ALV-induced erythroblastosis: novel RNA processing and

A:Reference number: A00643; MIMD:85228222; PMID:258784

A:Accession: A00643

A:Molecule type: mRNA

A:Residues: 585-1223 <NTL>

A:Cross-references: GB:M10066

C:Genetics:

A:Gene: erbB

C:Superfamily: epidermal growth factor receptor; protein kinase homology

C:Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; growth fact

specific protein kinase

F:1-30/Domain: signal sequence #status predicted <SIG>

F:31-123/Domain: epidermal growth factor receptor #status predicted <MAT>

F:31-654/Domain: extracellular #status predicted <EX>

F:81-307/Domain: EGF receptor extracellular domain repeat <BE1>

F:397-610/Domain: EGF receptor extracellular domain repeat <EE2>

F:655-677/Domain: transmembrane #status predicted <TM>

F:678-1223/Domain: intracellular #status predicted <INT>

F:719-984/Domain: protein kinase homology <KIN>

F:727-735/Region: protein kinase ATP-binding motif

F:136,202,280,361,422,575,580,615,635/Binding site: carbohydrate (Thr) (covalent)

F:192,650/Binding site: carbohydrate (Ser) (covalent) #status predicted

F:687/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predite

F:754/Active site: Lys #status predicted

F:1100,1183,1208/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #

Query Match 38.2%; Score 1509.5; DB 1; Length 1223;  
 Best Local Similarity 44.2%; Pred. No. 7.2e-93;  
 Matches 286; Conservative 103; Mismatches 235; Indels 23; Gaps 9;

QY 8 RWGILLALLPFGAA-----STOYCTGTDMKRLRPASPEHLDMLRLHYGCCOYVGGNLE 61  
 Db 13 RGAIVELVLLILGLVACSVEERKVCQGTNNKRLTQLGHEVDFHTSLQRYNNCEVLSNLE 72  
 QY 62 LITVPNASSLFDIOEVQGYVLIANQVROYQLRLRYRTQLEFEDNALVALVDNGDPLNNTTP 121  
 Db 73 ITVEHNRDLFTKTIQEVAGYVLIANTVERIPLENQILRGNALENTYALAILSN----- 124  
 QY 122 PLNNTPTVGTASPGRLRELQRLSTLEILKGVLIQRNPQLCYODTILMKDI-----FHKNNOLATLI 181  
 Db 133 -MKNTQ-----GLRELPMKRLSELNCGVKSINNPKICNMDYTMNDIIDTSRK-PL 182  
 QY 182 TLID-TNRSRACHPCSPKCSRCMGESSEDCQSLRTVCAGCA-RCKGPLPDDCCHEQ 239  
 Db 183 TVLDFASNLSSPCPKCHPNCNEDHCWGAEONCOTITFVICAQOCSGRKGRKVPSSDCCHQ 242  
 QY 240 CAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESMPNPEGRYFGASCYACPYNY 299  
 Db 243 CAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESMPNPEGRYFGASCYACPYNY 299  
 QY 300 PYNLTSTVGSCTIVCPLIHQNEVTAEDGTQRCERKCSKPCARVCYGLGMEHLREYRAVTSANIOE 363  
 Db 300 PYNLTSTVGSCTIVCPLIHQNEVTAEDGTQRCERKCSKPCARVCYGLGMEHLREYRAVTSANIOE 363

```

Db      303 PHNYVVTBHGSCVBRSCNDTYEV - EENGVRKCKCKDDGLCSKVCNIGIGELKGIILSTNAT 361
OY      360 NIQFAGACKKJFGSLAFJPRESEFDGPASNTAPLQREQLOVFETLEETGYLISAMPDSL 419
Db      362 NIDSEKNTCTKINGDVSIIPVAFJGDAFTKTLPLDERKIDVFRTYKEISOFELIQAMPDWA 421
OY      420 PDLVSFOULQVYRGILNLNGAVSLTLOGIGISWGLRSTRIRELSGSLALIHNTHTLCFVHT 479
Db      422 TDLVAFENLEIRGHTKOHGOVSLAVNLIKISLQSLRSLKETISDDIDIAIMKKNKCYADT 481
OY      480 VPMIDLFENRPNQALHTHTANPREDECVGEGELACHOCARGHCMPGPOTVCNCSOPLRQOE 539
Db      482 MNMRSLFNTQSGKXTIIONRKNNDCTADRHVCDPLCSVYGCMGPRPFHCFSCRFSTRQKE 541
OY      540 CVEBCRVLOGLPREYVANHRLCPHRECPONG--SVYTCFGRPADQCVCACHYKDPFC 596
Db      542 CVKOCNIIIOGPREEREDRSKCLPCHSECLVONSTAYNTTCSOPGPDHMKCAHFDGBHC 601
OY      597 VARCPSGVPRLSVMPITKPFDEBEGACQPCPINCNSCVDDDDKCCP 643
Db      602 VKACAGVGLGENDTL-VMKYADANAVQDLCHNCNCRGCKGPGLESCP 647

```

## RESULT 8

A47253  
epidermal growth factor receptor, HER4 - human  
C:Species: Homo sapiens (man)  
C:Date: 22-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 18-Jun-1999  
C:Accession: A47253  
R:PiDownman, G.D.; Culusconu, J.M.; Whitney, G.S.; Green, J.M.; Carlton, G.W.; Foy, L.; N  
Proc. Natl. Acad. Sci. U.S.A. 90, 1746-1750, 1993  
A:Title: Ligand-specific activation of HER4/Pl80erbB4, a fourth member of the epidermal  
A:Reference number: A47253; MUID:93189574; PMID:8383326  
A:Accession: A47253  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: nucleic acid  
A:Residues: 1-1308 <P>LO  
A:Cross-references: GB:U07868; NID:q337359; PIDN:AAB59446.1; PID:q337360  
A:Ntce: sequence extracted from NCBI backbone (NCBI:126842)  
C:Superfamily: epidermal growth factor receptor; protein kinase homology  
C:Keywords: ATP; growth factor receptor  
E:710-981/Domain: protein kinase homology <KIN>  
E:724-732/Region: protein kinase ATP-binding motif

Query Match	36.7%	Score 1451	DB 2	Length 1308
Best Local Similarity	43.58%	Pred. No. 6.2e-89		
Matches 275; Conservative	99;	Mismatches 238;	Indels 20;	Gaps 8;

[illegible]

```

OY 364 PAAKKGSLALPEESDGPASNTAPLOEYLETTELGYLIAMPSLDLS 423
    | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 336 FINTCKINGNLLELTGTGHSPYNAIEADIEKLNWFETAREITGELINMSPPNMTDES 415
    | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 424 VFONLOVIRGRILHNGAYSLTLOGIGISWGLRSILRELGSGLALIHNTHCYVHPMD 483
    | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 416 VESNLVTIGGRVLSGSLLLILKQOQJTSLOFQISKEISAGNITYITDNSNLCYHTINT 475
    | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 484 OLFRRPHALLHTANRPDECEVAGSLACHOLCAAGHCNMGSPQCVACSOFLGQBEVEE 543
    | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 476 TLESTINORIYIRDRKAKENCTAGSMVCNHLCSDDGCMGEPDCLSKRRFSKRCTIES 535
    | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 544 CRVLOGLEREVYNARHCLPCHRECP-ONGSVTFEGEADQCVACAHYKDPFCVACPS 602
    | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 536 CNLYGEEFREFENGSIYECSDPOCEKMDGLTLCGHGPDNCTGCSHFKDGPNCEKCPD 595
    | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 603 GVKPDLSTYMPKRPDEGACORPINCOTHS 634
    | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 596 GIGGANSF--LFTYADPRDRECHCPAPNCTQCC 625
    | | | | | | | | | | | | | | | | | | | | | | | | | |

```

## RESULT 9

A36223  
kinase-related transforming protein (erbB3) (EC 2.7.1.-) precursor - human  
C.Species: Homo sapiens (man)  
C.Date: 04-Oct-1991 #sequence\_revision 13-Jan-1993 #text\_change 17-Nov-2000  
C.Accession: A36223; 159164  
R.Kraus, M.H.; Issing, W.; Miki, T.; Popescu, N.C.; Aaronson, S.A.  
Proc. Natl. Acad. Sci. U.S.A. 86, 9193-9197, 1989  
A.Title: Isolation and characterization of ERBB3, a third member of the ERBB/epiderma  
A.Reference number: A36223; MUID:90083234; PMID:2687875  
A.Accession: A36223  
A.Status: preliminary  
A.Molecule type: mRNA  
A.Residues: 1-1342 <KRA>  
A.Cross-references: GB:M29366  
Proc. Natl. Acad. Sci. U.S.A. 87, 4905-4909, 1990  
A.Title: Molecular cloning and expression of another epidermal growth factor receptor  
A.Reference number: 159164; MUID:90311312; PMID:2164210  
A.Accession: 159164  
A.Status: preliminary; translated from GB/EMBL/DBJ  
A.Molecule type: mRNA  
A.Residues: 1-559, 'G', 561-957, 'F', 959-1063, 'G', 1065-1342 <RES>  
A.Cross-references: GB:M34309; NID:g183990; PIDN:AAA35979.1; PID:g306841  
C.Genetics:  
A.Gene: GDB:ERBB3; HER3  
A.Cross-references: GDB:119880; OMIM:190151  
A.Map position: 12q13-12q13  
C.Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom  
C.Keywords: ATP, phosphotransferase  
C.7107-912/Domain: protein kinase homology <KIN>  
C.7115-723/Region: protein kinase ATP-binding motif

Query Match	36.3%	Score 1435	DB 2	Length 1342
Best Local Similarity	43.6%	Pred. No. 7.5e-88		
Matches 278	Conservative 92	Mismatches 244	Indels 24	Gaps 12

[illegible]

OY	247	PKHSHCLACIHPHNHSGIOELHCPALVYTNITDFEESMPNDEGKYTTGASGVACRYNXYLT	306
Db	238	PODITCFACRHHNDGAGCPRCPDPLVYTKLFEPLERPHNTYQYGGVACVASCPIHNF-V	296
OY	307	DVGSCTIYCPALHNOVTAEDGTORCEKSCPCARVCYGIEMHELRVAVTASANTOEPAG	366
Db	297	DOTSCVRACPRDKMEVD-KNGIKMCEPCGGCLCPKACEEGSS--SRFQTVDSNIDGCVN	353
OY	367	CKKIFGSLAFDESDGDPASNTAPLQBPQLOVFETLEETGYLYXISAMPDSLPLSVFO	426
Db	354	CTKIGINDLFTLTGGNDPWHKIPALDEPKLVFRTVREITGYLNIQSWPMHMFVFS	413
OY	427	NLOVIRGRILNHGAVS-LTLOGLGISTWGLRSLRLELGSGLALHNHTLCEPHNYTPMOQL	485
Db	414	NLTGGSLSLYRGFSLMKMLNVTSLGFRSLKJISAGRIYISANROLCYHNHSLMNTKV	473
OY	486	FNPQHALLHTA-NPREDCEVGEGLACHOLCARGCHMGWGPOTCVNCSOFLRGDCEVEBC	544
Db	474	LRGPFERERLDIKHNPRPRDCAVEGKYCDPLSSGGSCWBPQCLISCRNYSISGVCYTHC	533
OY	545	RYVLOGLPREYVNAHRCLEPNECOPQNGSVTCFGRADQCVACAHAHYKDRPFVCAPSGV	604
Db	534	NFLNPEPRFEAAHEACFSCHPEQCMEGATNGSGSDPICAQCAHFRDGRPHCVSSCPHGV	593
OY	605	KPDLSYMPIMKPRDEGACORPCPTICTHSCVDYLDKGC	642
Db	594	LG--AKGRPLYKPDVONCECRPCHEMCTOGCKRGKPELDDC	629

RESULT 10  
JC4387  
epidermal growth factor receptor homolog precursor - rat  
N:Alternate names: ErbB3 protein; HER3 protein  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 17-Jan-1996 #sequence\_revision 19-Apr-1996 #text\_change 13-Nov-1998  
C:Accession: JC4387  
R:Hellier, N.J.; Kim, H.H.; Greaves, C.H.; Sierke, S.L.; Koland, J.G.  
Gene 165, 279-284, 1995  
A:Title: Cloning of the rat ErbB3 cDNA and characterization of the recombinant protein.  
A:Reference number: JC4387; MUID:96096535; PMID:8522190  
A:Accession: JC4387  
A:Molecule type: mRNA  
A:Residues: 1-1339 <HEL>  
A:Cross-references: GB:U29339; NID:g915389; PID:g915390  
A:Experimental source: liver  
A:Note: The authors translated the codon AAC for residue 369 as Thr and GTT for residue  
C:Comment: This protein is a functional heregulin receptor that transduces signals to th  
C:Genetics:  
A:Gene: ErbB3  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo  
C:Keywords: ATP; growth factor receptor; liver; phosphoprotein; transmembrane protein  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-1339/Product: epidermal growth factor homolog #status predicted <MNT>  
F:640-659/Domain: transmembrane #status predicted <TMM>  
F:705-970/Domain: protein kinase homology <KIN>  
F:713-721/Region: protein kinase ATP-binding motif  
F:939,1051,1156,1194,1196,1219,1257,1259,1273,1286,1325/Binding site: phosphate (Tyr) (G

Query Match	34.2%	Score	1351.5	DB 2	Length	1339			
Best Local Similarity	42.2%	Pred. No.	2.8e-82						
Matches	269	Conservative	87	Mismatches	253	Indels	29	Gaps	13

[illegible]

```

Db 121 ---LayNt-----NSSHAlRQlKfTQlTqEtlISGvYIEkNKlChMDtIDMRdYRv--- 170
Oy 180 AlTlIDrRSrAChPCsMcGSRKwGSESSDcSlRtYvAGcC-ARckPlPTDcChE 238
Db 171 GAElYKKNNGANcPCpEcKvKcG-RcMGpGRDcClTlKtTlADpCmNGRcGpPmNcOCHD 229
Oy 239 oCAAcGtSPKkSDcLAcLhFmHNSGlcELhCPAlYtYmTDEESMpNEGRtFGaScvTA 298
Db 230 EcAGGcSSpRtDlDcFACaRfNDScAcYpRCePlYvNKlTfQLEpNRhTtKYqYcGvCvAS 289
Oy 299 cPYnLstDvGScTlYvCPlhNoEvTAEDGtQRcEKcSKpCARvcYGLmEHLEvRAvTs 358
Db 290 cPnhEv-VDOfEcYvACpPDkMEvD-KHGLMcPCcGGLcPKAcEGtGSG--sRYQvDVS 345
Oy 359 ANIOEcAcKkIFESlAFtPESFDGpDASnAPtAPDpDlOvFEtLEtIGtYtSAMPDS 418
Db 346 SNIDeFvAcTtIlGNDPltIGlVNDvPmKtPAlDpEKlNfYvREItGTlNtQSMpPh 405
Oy 419 LpDlSVFONtOVInGRtILhNGaYS-tLQlGtISmLgRSrLEtGSLAtlIHnHtLcFv 477
Db 406 MhNFsVFNlTtIGRSrLYNGfSLtIMKtNlNvTSlGfRSrLEtISAGrYtISANQlCYH 465
Oy 478 HtVpPDOLeFRmPhOAlLHTA-nRPEDeCVGSGlACHOLcAGHcWmGgPpQcVvCSQfLr 536
Db 466 hSLmNtRtLlRgPSErERtIKYDRPlGGEclAGKAcVDPtCSSGGGcWGaPQcScsRnYSr 525
Oy 537 GoEcEeCRvLQGLpREvYvNARhCLpChPEcOPONGSVtGFGpPDeCvACaHkDpPfc 596
Db 526 EGvCYtHcNFLOgEPREvFvHAcOfcSCHPEcLtpMEGtStYNGSSDcACARhARpDgPhC 585
Oy 597 vARcSSgYKpDlStYmPlMKfPDEGaCOPcPcIntHtGc 634
Db 586 vNScPhGtLG--AKGpLYKtPDaONECRpChENtQcG 621

```

RESULT 11  
S06142  
protein-tyrosine kinase (EC 2.7.1.112) mrk-v precursor - southern platyfish  
N:Alternate names: epidermal growth factor receptor homolog; kinase-related transform  
C:Species: Xiphophorus maculatus (southern platyfish)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 18-Feb-2000  
A:Accession: S06142; S13809  
R:Witbrodt, J.; Adam, D.; Malitschek, B.; Maeueller, W.; Raulf, F.; Telling, A.; Robe  
Nature 341, 415-421, 1999  
A:Title: Novel putative receptor tyrosine kinase encoded by the melanoma-inducing Tu  
A:Reference number: S06142; MUID:90015140; PMID:2797166  
A:Accession: S06142  
A:Molecule type: DNA  
A:Residues: 1-1166 <MIT>  
A:Cross-references: EMBL:X16891; NID:g65290; PIDN:CMA34770.1; PID:g65291  
R:Adam, D.; Maeueller, W.; Scharlt, M.  
Oncogene 6, 73-80, 1991  
A:Title: Transcriptional activation of the melanoma inducing xmrk oncogene in xiphoph  
A:Reference number: S13807; MUID:91125882; PMID:1846957  
A:Accession: S13809  
A:Status: preliminary; translation not shown  
A:Molecule type: DNA  
A:Residues: 821-1025, 'N', 1027-1098, 'A', 1100-1166 <ADA>  
A:Cross-references: EMBL:X56319; NID:g65284; PIDN:CMA39763.1; PID:g65285  
C:Genetics:

A:Gene: mirk  
A:Map position: Y  
A:Introns: 872/3: 898/1: 947/1: 979/3: 1025/3: 1056/1  
C:Superfamily: epidermal growth factor receptor; protein kinase homology  
C:Keywords: ATP; growth factor receptor; phosphotransferase; transmembrane protein; t  
F:1-75/Domain: signal sequence #status predicted <SIG>  
F:26-1166/Product: kinase-related transforming protein (Tu) #status predicted <KAT>  
F:707-972/Domain: protein kinase homology <KIN>  
F:715-723/Region: protein kinase ATP-binding motif

Query Match 33.2%; Score 1313.5; DB 1; Length 1166;  
Best Local Similarity 41.6%; Pred. No. 8.3e-08;  
Matches 268; Conservative 93; Mismatches 259; Indels 25; Gaps 12;





A:Residues: 52-1374 <MIL>  
A:Cross-references: EMBL:Z70038; PIDN:CAA93882.1; GSPDB:GN00020; CESP:ZK1067.1  
A:Experimental source: clone ZK1067  
C:Genetics:  
A:Gene: let-23; CESP:ZK1067.1  
A:Map position: 2  
A:Introns: 44/1; 51/1; 90/1; 112/3; 165/3; 217/1; 290/1; 379/1; 418/1; 448/2; 565/1; 608/1  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
C:Keywords: ATP; phosphotransferase; transmembrane protein; tyrosine-specific protein kinase  
F:1-28/Domain: signal sequence #status predicted <Sig>  
F:29-1374/Product: protein-tyrosine kinase let-23 #status predicted <Mat>  
F:934-1199/Domain: protein kinase homology <Kin>  
F:942-950/Region: protein kinase ATP-binding motif

Query Match 18.2% Score 718.5; DB 2; Length 1374;  
Best Local Similarity 28.0%; Pred. No. 5.1e-40;  
Matches 191; Conservative 107; Mismatches 256; Indels 129; Gaps 28;

QY 25 VCTGDMKRLRPASPEHRLDMRLHLYGQYVGNLELYPTN----- 68  
Db 90 LCSGTNGISRYGTGNI-LEDEFTMYRCRRYVGLLEITWIANEIKKRESTNSTVDPK 148  
QY 69 -----ASLSFLQDIQEVGYVLAHQVROVPLQRLIRVGTQLFEDNYALAVLNGDP 122  
Db 149 NEDSPLKSINFEDMLEIRGSLIYRANIKISPEPLRVIYGEVFDHN-ALYIHKNDK- 206  
QY 123 LNNTPVVGASPGGRLREQLRSLEILKGVLIQNRPOLCY-QDTILMKDIFHKNNQLAL 181  
Db 207 -----VHEVVMRELRYIRNGSVYTIQDNPKCYIGDKIDMKELLYDPD--VQ 250  
QY 182 TLIDTNRSRACH-----PCSPMKGSRCWGESEDCOSLTRVYACAGCARC---KGPL 231  
Db 251 KVEETNSHQHCYQNGKSMARKCHESC-NDKMGSGDNDQCQRYRYSVCPSQCFYSNSTS 309  
QY 232 PTDCHEQCAAGCTGPKHSDCLACLFHNSGICELCPALVTYNTDTPESMPNPEGRTF 291  
Db 310 SYECDSACLGCTGHGPKKNCIACSKYELDGICIEPTCPSRKIFNHKTGRIVFNPGRYON 369  
QY 292 GASCVTACPPNYL-STDVGSCTLWC-PLHNOEVTAEDGTORCEK-SKPCARVCYGLGME 348  
Db 370 GNHCYKCEPPELLIENDV--CVRHCSGDGHAYDAIKD--VRECEKCRSSSCPKICTVDG-- 423  
QY 349 HIREYRAVTSANIOEFAGCKKIFGSLAFLPESFDGDPASNTAPLOPELOVFEETLEITG 408  
Db 424 HL-----TNETLKNLECEGEQIDGHL-IIEHAF-----TYEQKYLEYKAIYSE 465  
QY 409 YLYISAMPDSLPLDSVFQNLQVINGRILIHNGAYSITL-QGLGISWLGRLSRLGSGLAL 467  
Db 466 YITIV--QQNFYDLKFLKNIQIIEGRKLIHNVMAIAIYQCDLLELSLSLKLITGAVL 523  
QY 468 IHHHTHLCEFTVTPWDOLF-----RNPHQALLHTANRPDECEVSGSLACHQICARGHCWGP 523  
Db 524 IMKNHRLCYVASKIDWSSITTSKGDKNRPSLIAIAENRDSKLCETEDQRYCDKNCKRGK 583  
QY 524 GPFOCVNCSQFLRGQCEVEECVRLQGLPREYVNAHNCILPCHECOPONGSVTCFGEADQ 583  
Db 584 EPEDCLECKTKYSVGTCEYKCDT-KGFLANQTSMK-CERCSPECE-----TCNGLGELD 635  
QY 584 CVACAH-----YKDPFCVAKCPSGVKPDLSTYMPIWKFPEDEGACQPCPINCTHSCVD 636  
Db 636 CLTCRHKTKLYNSDFGNRMCEVHDC-----PVSHFPTQKNVCEKCHPTCY----- 679  
QY 637 LDDKGCPRAEGRASPLTSQNEIDLG 659  
Db 680 --DNKC-----TGPDNSNIDG 691

Search completed: January 13, 2003, 14:49:20  
Job time : 23.4617 secs





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2003, 14:47:00 : Search time 10.7057 Seconds  
(without alignments)  
1665.428 Million cell updates/sec

Title: US-09-854-356-6

Perfect score: 5078  
Sequence: 1 METALCRMGILLALLPPGA.....RFKGPTRAEINPEYGLDYPV 919

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 118974 seqs, 19401057 residues

Total number of hits satisfying chosen parameters: 118974

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpa/US08\_NEW\_PUB\_PEP:\*
- 2: /cgn2\_6/ptodata/1/pubpa/PCT\_NEW\_PUB\_PEP:\*
- 3: /cgn2\_6/ptodata/1/pubpa/US06\_NEW\_PUB\_PEP:\*
- 4: /cgn2\_6/ptodata/1/pubpa/US06\_PUBCOMB\_PEP:\*
- 5: /cgn2\_6/ptodata/1/pubpa/US07\_NEW\_PUB\_PEP:\*
- 6: /cgn2\_6/ptodata/1/pubpa/US07\_PUBCOMB\_PEP:\*
- 7: /cgn2\_6/ptodata/1/pubpa/PCTUS\_PUBCOMB\_PEP:\*
- 8: /cgn2\_6/ptodata/1/pubpa/US08\_PUBCOMB\_PEP:\*
- 9: /cgn2\_6/ptodata/1/pubpa/US09\_NEW\_PUB\_PEP:\*
- 10: /cgn2\_6/ptodata/1/pubpa/US09\_PUBCOMB\_PEP:\*
- 11: /cgn2\_6/ptodata/1/pubpa/US10\_NEW\_PUB\_PEP:\*
- 12: /cgn2\_6/ptodata/1/pubpa/US10\_PUBCOMB\_PEP:\*
- 13: /cgn2\_6/ptodata/1/pubpa/US60\_NEW\_PUB\_PEP:\*
- 14: /cgn2\_6/ptodata/1/pubpa/US60\_PUBCOMB\_PEP:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5078	100.0	919	9 US-09-854-356-6	Sequence 6, Appli
2	4900	96.5	1255	9 US-09-854-356-1	Sequence 1, Appli
3	4900	96.5	1255	9 US-09-930-125-2	Sequence 2, Appli
4	4892	96.3	1255	10 US-09-811-123-9	Sequence 9, Appli
5	4892	96.3	1255	10 US-09-811-115-3	Sequence 3, Appli
6	4843	95.4	1255	9 US-09-769-508-2	Sequence 2, Appli
7	4138.5	81.5	1256	9 US-09-854-356-2	Sequence 2, Appli
8	4134	81.4	1260	9 US-09-870-759-118	Sequence 118, App
9	4125.5	81.2	1256	9 US-09-854-356-14	Sequence 14, Appli
10	3954	77.9	712	9 US-09-854-356-7	Sequence 7, Appli
11	3628	71.4	653	9 US-09-854-356-3	Sequence 3, Appli
12	3590	70.7	645	10 US-09-921-161-1	Sequence 1, Appli
13	3110.5	61.3	654	9 US-09-854-356-8	Sequence 8, Appli
14	2585	50.9	564	10 US-09-821-883-3	Sequence 3, Appli
15	2585	50.9	697	10 US-09-821-883-4	Sequence 4, Appli
16	2573.5	50.7	555	10 US-09-821-883-1	Sequence 1, Appli
17	2573.5	50.7	690	10 US-09-821-883-2	Sequence 2, Appli
18	1645	32.4	1210	10 US-09-725-433-2	Sequence 2, Appli
19	1608	31.7	479	10 US-09-821-883-5	Sequence 5, Appli

20	1587	31.3	289	10 US-09-821-883-23	Sequence 23, Appli
21	1528.5	30.1	1308	10 US-09-940-101-2	Sequence 2, Appli
22	1474	29.0	583	9 US-09-930-125-9	Sequence 9, Appli
23	1474	29.0	587	9 US-09-930-125-8	Sequence 8, Appli
24	1474	29.0	589	9 US-09-930-125-10	Sequence 10, Appli
25	1474	29.0	600	9 US-09-930-125-11	Sequence 11, Appli
26	1450	28.6	266	9 US-09-854-356-4	Sequence 4, Appli
27	1447	28.5	615	10 US-09-940-101-4	Sequence 4, Appli
28	1177	23.2	217	10 US-09-821-883-25	Sequence 25, Appli
29	1177	23.2	397	10 US-09-821-883-27	Sequence 27, Appli
30	1177	23.2	1179	10 US-09-821-883-29	Sequence 29, Appli
31	1028.5	20.3	1438	10 US-09-867-521-2	Sequence 2, Appli
32	647.5	12.8	475	10 US-09-925-301-1232	Sequence 1232, Ap
33	326	6.4	59	9 US-09-854-356-5	Sequence 5, Appli
34	311.5	6.1	1367	9 US-09-870-759-120	Sequence 120, App
35	281	5.7	370	10 US-09-205-658-104	Sequence 104, App
36	267.5	5.3	1724	10 US-09-844-353A-104	Sequence 104, App
37	267.5	5.3	1724	10 US-09-205-658-12	Sequence 12, Appli
38	267.5	5.3	1724	10 US-09-844-353A-12	Sequence 12, Appli
39	250	4.9	366	10 US-09-205-658-103	Sequence 103, App
40	250	4.9	366	10 US-09-844-353A-103	Sequence 103, App
41	223.5	4.4	383	10 US-09-205-658-105	Sequence 105, App
42	223.5	4.4	383	10 US-09-844-353A-105	Sequence 105, App
43	212	4.2	381	10 US-09-205-658-106	Sequence 106, App
44	212	4.2	381	10 US-09-844-353A-106	Sequence 106, App
45	199	3.9	830	9 US-09-870-759-140	Sequence 140, App

#### ALIGNMENTS

```
RESULT 1
US-09-854-356-6
; Sequence 6, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Cheysen, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 919
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: fusion protein
; OTHER INFORMATION: Of ECD and PD of human HER-2/neu
US-09-854-356-6

Query Match          100.0%; Score 5078; DB 9; Length 919;
Best Local Similarity 100.0%; Pred. No. 2e-285;
Matches 919; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 METALCRMGILLALLPPGAASVCTGTGDMKRLRPASETHIDMLRHLYGCGVQVGM 60
DB 1 METALCRMGILLALLPPGAASVCTGTGDMKRLRPASETHIDMLRHLYGCGVQVGM 60
QY 61 ELTYLPTNASLSLTDIQYQGVVLIHNOVQVPLQRLRIYNGTGLFEDNVALAVLDNG 120
DB 61 ELTYLPTNASLSLTDIQYQGVVLIHNOVQVPLQRLRIYNGTGLFEDNVALAVLDNG 120
QY 121 DPLNNTPTVGTGSPGLRFLQSLFEILKGVGLIORNPOLCQDPTITLKKDFHKNNOIA 180
DB 121 DPLNNTPTVGTGSPGLRFLQSLFEILKGVGLIORNPOLCQDPTITLKKDFHKNNOIA 180
```

Db 121 DPLNNTPTVTGASPGELRELOLRSLTELLKGVLLIQRNPDLCYDPTLIMKDIIFHKNNOLA 180  
QY 181 LTLIDTNRSRACHPCSPCKSRCKGSESSDQSLTRTVCGAGCARCKGPLPTDCHQC 240  
Db 181 LTLIDTNRSRACHPCSPCKSRCKGSESSDQSLTRTVCGAGCARCKGPLPTDCHQC 240  
QY 241 AAGCTGPKHSDCLACLFHNHSGICELHCPALVTYNTDTFESMPNDEGRTTGASCVTAC 300  
Db 241 AAGCTGPKHSDCLACLFHNHSGICELHCPALVTYNTDTFESMPNDEGRTTGASCVTAC 300  
QY 301 YNYLSTDVSGCTVCPPLNNOEYTAEDGTQRCCKSKPCARVCYGLGMHLEBRAVATSAN 360  
Db 301 YNYLSTDVSGCTVCPPLNNOEYTAEDGTQRCCKSKPCARVCYGLGMHLEBRAVATSAN 360  
QY 361 IOEAGCKKIRGSLAFLEPSEFDGDPASNTAPLOPEQLOVFETLEITGYLYISAMPDLP 420  
Db 361 IOEAGCKKIRGSLAFLEPSEFDGDPASNTAPLOPEQLOVFETLEITGYLYISAMPDLP 420  
QY 421 DLSVFONLOVIRGRILHNGAYSLTLOGIGISMLGRLSRLRELSGSLALIHNTHLCEFTV 480  
Db 421 DLSVFONLOVIRGRILHNGAYSLTLOGIGISMLGRLSRLRELSGSLALIHNTHLCEFTV 480  
QY 481 PWDOLFRRPHQALLHTANRPDECEYGEGLACHQOLCARHCHGPGPTQCVCNCSQFLRGEC 540  
Db 481 PWDOLFRRPHQALLHTANRPDECEYGEGLACHQOLCARHCHGPGPTQCVCNCSQFLRGEC 540  
QY 541 VEECRVLOGLPREYVYNAHCHLPCHECOPONGSVTCFGEADOCYACAHYKDRPPCVAR 600  
Db 541 VEECRVLOGLPREYVYNAHCHLPCHECOPONGSVTCFGEADOCYACAHYKDRPPCVAR 600  
QY 601 PSYKVPDLSTYMPIMKFPDEEGACQPCPINCTHSCVDLDKCPABORASPLTSQNEIDGP 660  
Db 601 PSYKVPDLSTYMPIMKFPDEEGACQPCPINCTHSCVDLDKCPABORASPLTSQNEIDGP 660  
QY 661 ASPDSTFYRSLDEDDMDGLVDAEYLVPOGFFCPDPAAGAGMHHRRSSSTRSGG 720  
Db 661 ASPDSTFYRSLDEDDMDGLVDAEYLVPOGFFCPDPAAGAGMHHRRSSSTRSGG 720  
QY 721 GDLTLGLEPSESEEARSPPLAPSEAGSDVPFDGLGMGAAGKQSLPTHDPSLQYSSDP 780  
Db 721 GDLTLGLEPSESEEARSPPLAPSEAGSDVPFDGLGMGAAGKQSLPTHDPSLQYSSDP 780  
QY 781 TVPLPSEIDGYVAPLTCSPQPEYVNOQPVPRQPPSPREGPLPARPACATLEPXTLSPG 840  
Db 781 TVPLPSEIDGYVAPLTCSPQPEYVNOQPVPRQPPSPREGPLPARPACATLEPXTLSPG 840  
QY 841 KNGVYKDVFAFGAVENNEYLTPOGGAPOPHPPAPSPADNLYYMDODDERGAPST 900  
Db 841 KNGVYKDVFAFGAVENNEYLTPOGGAPOPHPPAPSPADNLYYMDODDERGAPST 900  
QY 901 FKGTPTAENPEYLGIDVPV 919  
Db 901 FKGTPTAENPEYLGIDVPV 919

RESULT 2  
US-09-854-356-1

; Sequence 1, Application US/09854356  
; Patent No. US2002017567A1  
; GENERAL INFORMATION:  
; APPLICANT: Cheever, Martin A.  
; APPLICANT: Gheysen, Dirk  
; APPLICANT: Corixa Corporation  
; APPLICANT: SmithKline Beecham Biologicals S. A.  
; TITLE OF INVENTION: HER-2/neu Fusion Proteins  
; FILE REFERENCE: 014058-009810PC  
; CURRENT APPLICATION NUMBER: US/09/854,356  
; CURRENT FILING DATE: 2001-05-09  
; PRIOR APPLICATION NUMBER: US 09/493,480  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: US 60/117,976  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 26

; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 1255  
; TYPE: PR  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: human HER-2/neu protein  
; NAME/KEY: DOMAIN  
; LOCATION: (1)..(653)  
; OTHER INFORMATION: extracellular domain (ECD)  
; NAME/KEY: DOMAIN  
; LOCATION: (676)..(1255)  
; OTHER INFORMATION: intracellular domain (ICD)  
; NAME/KEY: DOMAIN  
; LOCATION: (990)..(1255)  
; OTHER INFORMATION: phosphorylation domain (PD)  
; NAME/KEY: DOMAIN  
; LOCATION: (1048)  
; OTHER INFORMATION: fragment of the phosphorylation domain, preferred  
; OTHER INFORMATION: portion (delta PD)  
US-09-854-356-1

Query Match 96.5%; Score 4900; DB 9; Length 1255;  
Best Local Similarity 73.2%; Pred. No. 5.3e-275;  
Matches 919; Conservative 0; Mismatches 0; Indels 336; Gaps 1;

QY 1 MELALCRWGLLALLPFGASTGVCTGDMKRLPASPETHLDMRLHYOGCCVVGNTL 60  
Db 1 MELALCRWGLLALLPFGASTGVCTGDMKRLPASPETHLDMRLHYOGCCVVGNTL 60  
QY 61 ELTYLPTNASLSFLQDIOEVGYVLIHNOYRQVPLQRLIRVGTQLEFEDNYALAVDNG 120  
Db 61 ELTYLPTNASLSFLQDIOEVGYVLIHNOYRQVPLQRLIRVGTQLEFEDNYALAVDNG 120  
QY 121 DPLNNTPTVTGASPGELRELOLRSLTELLKGVLLIQRNPDLCYDPTLIMKDIIFHKNNOLA 180  
Db 121 DPLNNTPTVTGASPGELRELOLRSLTELLKGVLLIQRNPDLCYDPTLIMKDIIFHKNNOLA 180  
QY 181 LTLIDTNRSRACHPCSPCKSRCKGSESSDQSLTRTVCGAGCARCKGPLPTDCHQC 240  
Db 181 LTLIDTNRSRACHPCSPCKSRCKGSESSDQSLTRTVCGAGCARCKGPLPTDCHQC 240  
QY 241 AAGCTGPKHSDCLACLFHNHSGICELHCPALVTYNTDTFESMPNDEGRTTGASCVTAC 300  
Db 241 AAGCTGPKHSDCLACLFHNHSGICELHCPALVTYNTDTFESMPNDEGRTTGASCVTAC 300  
QY 301 YNYLSTDVSGCTVCPPLNNOEYTAEDGTQRCCKSKPCARVCYGLGMHLEBRAVATSAN 360  
Db 301 YNYLSTDVSGCTVCPPLNNOEYTAEDGTQRCCKSKPCARVCYGLGMHLEBRAVATSAN 360  
QY 361 IOEAGCKKIRGSLAFLEPSEFDGDPASNTAPLOPEQLOVFETLEITGYLYISAMPDLP 420  
Db 361 IOEAGCKKIRGSLAFLEPSEFDGDPASNTAPLOPEQLOVFETLEITGYLYISAMPDLP 420  
QY 421 DLSVFONLOVIRGRILHNGAYSLTLOGIGISMLGRLSRLRELSGSLALIHNTHLCEFTV 480  
Db 421 DLSVFONLOVIRGRILHNGAYSLTLOGIGISMLGRLSRLRELSGSLALIHNTHLCEFTV 480  
QY 481 PWDOLFRRPHQALLHTANRPDECEYGEGLACHQOLCARHCHGPGPTQCVCNCSQFLRGEC 540  
Db 481 PWDOLFRRPHQALLHTANRPDECEYGEGLACHQOLCARHCHGPGPTQCVCNCSQFLRGEC 540  
QY 541 VEECRVLOGLPREYVYNAHCHLPCHECOPONGSVTCFGEADOCYACAHYKDRPPCVAR 600  
Db 541 VEECRVLOGLPREYVYNAHCHLPCHECOPONGSVTCFGEADOCYACAHYKDRPPCVAR 600  
QY 601 PSYKVPDLSTYMPIMKFPDEEGACQPCPINCTHSCVDLDKCPABORASPLTS 653  
Db 601 PSYKVPDLSTYMPIMKFPDEEGACQPCPINCTHSCVDLDKCPABORASPLTS 653  
QY 654 653  
Db 661 ILVVVLGVFGILIKRQOKIRKTYMRRLLQETELVEPLTPSGAMPNOAMRIKETEL 720

QY	654	-----	653
Db	721	RKYKVLGSGAFGTVYKGIWIPDGENYKIPVAIKIURENTSPKANKELIDEAYVYAGVGP	780
QY	654	-----	653
Db	781	YYSRLIGLTLSTVQLVTLQMLPGLCLLDHYRENRGRIGSODLIMCMQIANGMSTLEBYR	840
QY	654	-----	653
Db	841	LVIHRDLAARNVLSPKSNHKITDFGLARLLDDETHYADGSKVPIKMMALLESILRRFT	900
QY	654	-----	653
Db	901	HOSDVMYSYGVTWELMTFGAKPYDGIPAREIPDLLEKGERLPQPICTIDVYIMVYKCM	960
QY	654	-----QNDLCPASTLSTFYRSLLEDDMDGLDYA	684
Db	961	IDSECRPREFELVSEFSRMAKDPOREFVLIQNDLCPASTLSTFYRSLLEDDMDGLDYA	1020
QY	685	EEFLVPOOGFCFDPDPAAGAGVNHHRSSSTRSGGDLTLGLEPSEEDARSPPLAPSEG	744
Db	1021	EEFLVPOOGFCFCDPPAGAGGVNHHRSSSTRSGGDLTLGLEPSEEDARSPPLAPSEG	1080
QY	745	AGSDVFDGDLGMAAGLQSLPTHPDPSPLQKRSSEDPVPLSETDGYAARPLTCSPOREYV	804
Db	1081	AGSDVFDGDLGMAAGLQSLPTHPDPSPLQKRSSEDPVPLSETDGYAARPLTCSPOREYV	1140
QY	805	NORDVRRQRPSPREGFLPAARAGATLERPKTSLFGKNGVYKDYFAFGAVENBEYLTQ	864
Db	1141	NORDVRRQRPSPREGFLPAARAGATLERPKTSLFGKNGVYKDYFAFGAVENBEYLTQ	1200
QY	865	GGAARQNRPRARSPAFNDLIYWDODRRRGARSPSTFGKTPRAENBEYLTGLDVPY	919
Db	1201	GGAARQNRPRARSPAFNDLIYWDODRRRGARSPSTFGKTPRAENBEYLTGLDVPY	1255

### RESULT 3

```

US-9-930-125-2
US-9-930-125-2
SEQUENCE 2, Application US/09930125
Application NO. US20020193329A1
GENERAL INFORMATION:
APPLICANT: Hand-Zimmerman, Susan
APPLICANT: Cheever, Martin A.
APPLICANT: Foy, Teresa M.
APPLICANT: Lodes, Michael J.
APPLICANT: Ralos, Michael D.
APPLICANT: McNeill, Patricia D.
APPLICANT: Vedvick, Thomas S.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
OF HER-2/NEU-ASSOCIATED MALIGNANCIES
FILE REFERENCE: 210121.544
CURRENT APPLICATION NUMBER: US/09/930,125
CURRENT FILING DATE: 2001-08-14
NUMBER OF SEQ. ID NOS: 25
SOFTWARE: FastSeq for Windows Version 3.0
SEQ. ID NO. 2
LENGTH: 1255
TYPE: PRT
ORGANISM: Homo sapien
US-09-930-125-2
```

Query Match	96.5%	Score 4900;	DB 95;	Length 1255;
Best Local Similarity	73.2%	Pred. No. 5, 3e-275;		
Matches 919; Conservative	0;	Mismatches 0;	Indels 336;	Gaps 1;

QY 1 METALALCRGILLALLALPPGAASIVSTGCTGDMKRLRPASEPETHLDMRLHLYQGCGVQGNL 60  
1 METALALCRGILLALLALPPGAASIVSTGCTGDMKRLRPASEPETHLDMRLHLYQGCGVQGNL 60  
QY 61 ELTYLPNNAISLSPQDIQEVQGYVLLAHNOVROYPPQRLRTVKGDTLFFEDNRTALAVNDG 120

D	b	61	EDYLPN	NSLFIQD	IGEVQGY	VLI	HNQV	QVPLQ	R	RI	RIYRG	QLF	EDN	VAL	AV	LD	NS	120
Q	y	121	DLPLNT	PTV	TGAS	PGGL	RELQ	LR	SL	TE	IL	KG	VL	I	OR	N	PO	180
D	b	121	DLPLNT	PTV	TGAS	PGGL	RELQ	LR	SL	TE	IL	KG	VL	I	OR	N	PO	180
Q	y	181	LTLIDN	BR	A	H	PC	SP	MC	K	S	R	C	M	E	S	S	240
D	b	181	LTLIDN	BR	A	H	PC	SP	MC	K	S	R	C	M	E	S	S	240
Q	y	241	AAGCTG	PK	R	S	D	L	A	C	H	F	H	N	S	I	C	300
D	b	241	AAGCTG	PK	R	S	D	L	A	C	H	F	H	N	S	I	C	300
Q	y	301	YNYLST	D	W	S	C	T	L	Y	C	P	L	H	N	O	E	360
D	b	301	YNYLST	D	W	S	C	T	L	Y	C	P	L	H	N	O	E	360
Q	y	361	IOEFAG	CK	K	I	F	G	S	L	A	F	L	P	E	S	P	420
D	b	361	IOEFAG	CK	K	I	F	G	S	L	A	F	L	P	E	S	P	420
Q	y	421	DLSTVQ	N	O	L	O	V	I	G	R	I	L	H	N	G	A	480
D	b	421	DLSTVQ	N	O	L	O	V	I	G	R	I	L	H	N	G	A	480
Q	y	481	PMDOLE	FR	N	H	O	A	L	T	H	A	N	P	E	D	E	540
D	b	481	PMDOLE	FR	N	H	O	A	L	T	H	A	N	P	E	D	E	540
Q	y	541	VEBCRY	L	O	G	L	P	R	E	V	Y	N	A	R	H	C	600
D	b	541	VEBCRY	L	O	G	L	P	R	E	V	Y	N	A	R	H	C	600
Q	y	601	PEGV	P	D	L	S	T	Y	M	I	K	F	P	D	E	E	660
D	b	601	PEGV	P	D	L	S	T	Y	M	I	K	F	P	D	E	E	660
Q	y	654	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	653	
D	b	661	ILLYV	L	S	V	B	E	G	I	L	K	R	O	Q	K	I	720
Q	y	654	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	653	
D	b	721	RKYV	L	S	G	A	F	T	Y	K	G	I	W	I	P	D	780
Q	y	654	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	653	
D	b	781	YVSR	L	I	G	I	C	T	S	T	V	O	L	Y	T	O	840
Q	y	654	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	653	
D	b	841	LVHR	L	A	R	N	V	L	K	S	E	N	H	Y	K	I	900
Q	y	654	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	653	
D	b	901	HQSD	W	S	G	V	T	W	E	L	M	T	F	G	A	K	960
Q	y	654	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	653	
D	b	961	IDSE	C	R	P	R	E	L	V	S	E	P	S	R	M	A	1020
Q	y	685	EEY	L	V	P	O	O	G	E	F	C	P	D	A	P	A	1080
D	b	1021	EEY	L	V	P	O	O	G	E	F	C	P	D	A	P	A	1080
Q	y	745	AGSD	V	F	D	G	L	G	K	A	G	L	O	S	L	T	804
D	b	1081	AGSD	V	F	D	G	L	G	K	A	G	L	O	S	L	T	804
Q	y	805	NO	P	D	V	R	P	O	P	S	R	E	G	D	L	P	864
D	b	1141	NO	P	D	V	R	P	O	P	S							

QY 865 GGAAPQHPHPAPSPAFDNLXYWDODPPERKAGAPSTFKGPTAENPEYLGLDVPV 919  
 Db 1201 GGAAPQHPHPAPSPAFDNLXYWDODPPERKAGAPSTFKGPTAENPEYLGLDVPV 1255

RESULT 4  
 US-09-811-123-9  
 ; Sequence 9, Application US/09811123  
 ; Patent No. US20020001587A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sharon Erickson  
 ; APPLICANT: Ralph Schwall  
 ; APPLICANT: Mark Sliwowski  
 ; TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-ERBB  
 ; TITLE OF INVENTION: ANTIBODY-MAYTANSINOID CONJUGATES  
 ; FILE REFERENCE: GENENT 073A2  
 ; CURRENT APPLICATION NUMBER: US/09/811,123  
 ; CURRENT FILING DATE: 2001-03-16  
 ; PRIOR APPLICATION NUMBER: 60/238,327  
 ; PRIOR FILING DATE: 2000-10-05  
 ; PRIOR APPLICATION NUMBER: 09/602,530  
 ; PRIOR FILING DATE: 2000-06-23  
 ; NUMBER OF SEQ ID NOS: 11  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; ID NO 9  
 ; LENGTH: 1255  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-811-123-9

Query Match 96.3%; Score 4892; DB 10; Length 1255;  
 Best Local Similarity 73.1%; Pred. No. 1,5e-274;  
 Matches 918; Conservative 0; Mismatches 1; Indels 336; Gaps 1;

QY 1 METALALCRMGILLALPRGASVOTGCTMKLRPASPERHMLMRLHYGCGVGNL 60  
 Db 1 METALALCRMGILLALPRGASVOTGCTMKLRPASPERHMLMRLHYGCGVGNL 60  
 QY 61 ELTYLPTNASLFLQDIOEVGYVLIHNOYQVPLQRLIRVSTQLEFEDYALAVLDNG 120  
 Db 61 ELTYLPTNASLFLQDIOEVGYVLIHNOYQVPLQRLIRVSTQLEFEDYALAVLDNG 120  
 QY 121 DPLNNTPTVPGASPGGLREQLRSLTEILKGVLIQRNPOLCYODTILMKDIFKNNOLA 180  
 Db 121 DPLNNTPTVPGASPGGLREQLRSLTEILKGVLIQRNPOLCYODTILMKDIFKNNOLA 180  
 QY 181 LTLIDTRSRACHPCSPMGCSRCMGSSSEDCOSTRTVCGAGCARCKGRLPTCCHEQC 240  
 Db 181 LTLIDTRSRACHPCSPMGCSRCMGSSSEDCOSTRTVCGAGCARCKGRLPTCCHEQC 240  
 QY 241 AAGCTGKHSDDLACLFHNSGICELCPALVTYNTDFESMPNDEGRYTGASCVTACP 300  
 Db 241 AAGCTGKHSDDLACLFHNSGICELCPALVTYNTDFESMPNDEGRYTGASCVTACP 300  
 QY 301 YNLTSTVSGCTIVCPILHNOEVTAEDEGTQRCCKSKPCARVCYGLGMEHLREVRAYTSAN 360  
 Db 301 YNLTSTVSGCTIVCPILHNOEVTAEDEGTQRCCKSKPCARVCYGLGMEHLREVRAYTSAN 360  
 QY 361 IQERAGCKKTFGSLAFIPESFDGPRASNTAPLOPELOVFEETLEITGYLXISAMPDLP 420  
 Db 361 IQERAGCKKTFGSLAFIPESFDGPRASNTAPLOPELOVFEETLEITGYLXISAMPDLP 420  
 QY 421 DLSTFQMLQYIRGRILNNGAVSLTLOGISWLGRLSRLRETSGLALHNNTHLCEVHTV 480  
 Db 421 DLSTFQMLQYIRGRILNNGAVSLTLOGISWLGRLSRLRETSGLALHNNTHLCEVHTV 480  
 QY 481 PMDOLFRRNPQALHTANRPEDECVGBGLACHOLCARGHCWGPRTQCVNCSQFLRGDEC 540  
 Db 481 PMDOLFRRNPQALHTANRPEDECVGBGLACHOLCARGHCWGPRTQCVNCSQFLRGDEC 540  
 QY 541 VEECRVLQGLPREYVNAHCLPCHPECOPONGSVTCSGPEADQVACAHYDDPPPCVARG 600  
 Db 541 VEECRVLQGLPREYVNAHCLPCHPECOPONGSVTCSGPEADQVACAHYDDPPPCVARG 600

Db 541 VEECRVLQGLPREYVNAHCLPCHPECOPONGSVTCSGPEADQVACAHYDDPPPCVARG 600  
 QY 601 PSYKVPDLSTYMPIMKPEDEBACOPCPINCTHSCVDLDDKGCBAQERASPLTS----- 653  
 Db 601 PSYKVPDLSTYMPIMKPEDEBACOPCPINCTHSCVDLDDKGCBAQERASPLTSIVSAVVG 660  
 QY 654 ----- 653  
 Db 661 ILVVVLGVVFGILIKRRQOKIRKTYMRLLQETELVPLPSGAMPNOAMRLAETEL 720  
 QY 654 ----- 653  
 Db 721 RKVKVLGSGAFGYIKGIMWIPDGENVKIPVAIKVLRNTSPKANKELIDAIYVAGVGP 780  
 QY 654 ----- 653  
 Db 781 YSRLLGICLTSTVQLVQLMPYGCCLLDHYRENRGLSGODLLNMCQIAKMSYLEDVR 840  
 QY 654 ----- 653  
 Db 841 LVHRDLARNVLYKSPNHVKITDGLARLLDIDETEXHADGKVPYIKMMALESILRRFT 900  
 QY 654 ----- 653  
 Db 901 HQSDVMSGYVWELMTFGAKPYDGIPTAREIPDLLEKGERLPPPICTIDVYIMVKKCM 960  
 QY 654 -----ONEDIGRASPDLSTFYRSLLEDDMGDLVDA 664  
 Db 961 IDSECRPRRELIVSEFSRMAKDPQRFVYIIONEDIGRASPDLSTFYRSLLEDDMGDLVDA 1020  
 QY 685 EBYLVPOOGFPCPPRAGAGMWHHRSSSTRSGGDLTLGLEPSEEARSPPLASEG 744  
 Db 1021 EBYLVPOOGFPCPPRAGAGMWHHRSSSTRSGGDLTLGLEPSEEARSPPLASEG 1080  
 QY 745 AGSDVFDGDLGMAKAGLQSLPTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQREYV 804  
 Db 1081 AGSDVFDGDLGMAKAGLQSLPTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQREYV 1140  
 QY 805 NOPDVRPQPSPREGRLPAARPAAGATLERPKTUSPGKNGVYKDYFAFGAVENPEYLTPO 864  
 Db 1141 NOPDVRPQPSPREGRLPAARPAAGATLERPKTUSPGKNGVYKDYFAFGAVENPEYLTPO 1200  
 QY 865 GGAAPQHPHPAPSPAFDNLXYWDODPPERKAGAPSTFKGPTAENPEYLGLDVPV 919  
 Db 1201 GGAAPQHPHPAPSPAFDNLXYWDODPPERKAGAPSTFKGPTAENPEYLGLDVPV 1255

RESULT 5  
 US-09-811-115-3  
 ; Sequence 3, Application US/09811115  
 ; Patent No. US20020035736A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Erickson, Sharon  
 ; APPLICANT: Schwall, Ralph  
 ; APPLICANT: King, Kathleen  
 ; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL  
 ; FILE REFERENCE: GENENT 034A  
 ; CURRENT APPLICATION NUMBER: US/09/811,115  
 ; CURRENT FILING DATE: 2001-03-16  
 ; PRIOR APPLICATION NUMBER: 60/189,844  
 ; PRIOR FILING DATE: 2000-03-16  
 ; NUMBER OF SEQ ID NOS: 4  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 3  
 ; LENGTH: 1255  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-811-115-3

Query Match 96.3%; Score 4892; DB 10; Length 1255;  
 Best Local Similarity 73.1%; Pred. No. 1,5e-274;  
 Matches 918; Conservative 0; Mismatches 1; Indels 336; Gaps 1;

0Y	1	MEDALC8WGLLLALLPRGAASVQVCTGTMMKRLRLASRPTHLDMLRHLXQSCVYVQGM	60
0Y	1	MELALC8WGLLLALLPRGAASVQVCTGTMMKRLRLASRPTHLDMLRHLXQSCVYVQGM	60
0Y	61	ELTJLPTNASLSFJQDIQEOGVYLJAHNOVOPVLPORLRYVGTOLFEDNALAVLNG	120
0Y	61	ELTJLPTNASLSFJQDIQEOGVYLJAHNOVOPVLPORLRYVGTOLFEDNALAVLNG	120
0Y	121	DPLNNTPTVTCASPGCLREJOLSLREILKGVLIQRNPOLCYODTLMKDIFHKNNQJL	180
0Y	121	DPLNNTPTVTCASPGCLREJOLSLREILKGVLIQRNPOLCYODTLMKDIFHKNNQJL	180
0Y	181	LTLIDTNRSPACHCSPWCKSGSRCKWGBSSDDCSLRTVACAGGACRCKRLPTDCHEQC	240
0Y	181	LTLIDTNRSPACHCSPWCKSGSRCKWGBSSDDCSLRTVACAGGACRCKRLPTDCHEQC	240
0Y	241	AAGCGPRHSQCLCLFPHNSGTCIELCPALVYUINDTFESMNPBGRYTFGASCTYAP	300
0Y	241	AAGCGPRHSQCLCLFPHNSGTCIELCPALVYUINDTFESMNPBGRYTFGASCTYAP	300
0Y	301	YNYLSTFVSGCTVCPJLHNOEYTAEDOTCEKCSKPCARVCYGLMEHLREYRAVTSAN	360
0Y	301	YNYLSTFVSGCTVCPJLHNOEYTAEDOTCEKCSKPCARVCYGLMEHLREYRAVTSAN	360
0Y	361	IOEFAGCKITFGSLAFLPESFDGDPASNTAPLPORLOVEFTELEITGYLYISAMPDLP	420
0Y	361	IOEFAGCKITFGSLAFLPESFDGDPASNTAPLPORLOVEFTELEITGYLYISAMPDLP	420
0Y	421	DLSTFQNLQYRGRILHNGAYSLTLOGLGISWLGSLSLRELSGGLALHHNTHLCFVHTV	480
0Y	421	DLSTFQNLQYRGRILHNGAYSLTLOGLGISWLGSLSLRELSGGLALHHNTHLCFVHTV	480
0Y	481	PMDFLFNPHQALLHTANREDCEVGBGLACHOLCARGMCMGPRQVCNCSOFLRGQDC	540
0Y	481	PMDFLFNPHQALLHTANREDCEVGBGLACHOLCARGMCMGPRQVCNCSOFLRGQDC	540
0Y	541	VEBCRVLOGLPREXVYVNAHCLPCHPBCQDPONGSVTCFGEADQCVACAHYKDPRECVANC	600
0Y	541	VEBCRVLOGLPREXVYVNAHCLPCHPBCQDPONGSVTCFGEADQCVACAHYKDPRECVANC	600
0Y	601	PSGVRPDLSTMPYIMKPFDEGACQPCPINCTHSCVLDLCKGCPABEGRASPLTS-----	653
0Y	601	PSGVRPDLSTMPYIMKPFDEGACQPCPINCTHSCVLDLCKGCPABEGRASPLTS-----	653
0Y	654	ILLVVVLGVVFGILIKRRQOKIKKYMRRILQJTELVEPELTFSGAMPNOAQRILKETEL	720
0Y	654	ILLVVVLGVVFGILIKRRQOKIKKYMRRILQJTELVEPELTFSGAMPNOAQRILKETEL	720
0Y	721	RKVVLGSGAFGVYKGIWIPQENVKIPALIKVLERENTSPANKREILDEAVVMAGVSP	780
0Y	721	RKVVLGSGAFGVYKGIWIPQENVKIPALIKVLERENTSPANKREILDEAVVMAGVSP	780
0Y	781	VYSRLGICLSTVOLVTOQLMPYGLLDHYENRGRIGSODLNMCMQIANKMSYLEIDVR	840
0Y	781	VYSRLGICLSTVOLVTOQLMPYGLLDHYENRGRIGSODLNMCMQIANKMSYLEIDVR	840
0Y	841	LVLHDLARNAVLYKSPNHVKTTFPGGLARLLDIDETTYHADGCKVPIKMALESILRRRFT	900
0Y	841	LVLHDLARNAVLYKSPNHVKTTFPGGLARLLDIDETTYHADGCKVPIKMALESILRRRFT	900
0Y	901	HQDWSVXYGVVWELMTFGAKPYDGIIPAREIPDLLEKGERLPQPCITIDVYMIWKKCM	960
0Y	901	HQDWSVXYGVVWELMTFGAKPYDGIIPAREIPDLLEKGERLPQPCITIDVYMIWKKCM	960
0Y	961	IDSECRPRFRELVEFSRMAQDQRFVUJONEDLDGASPLDSTFYRSLLEDDMDKDLVDA	1020
0Y	961	IDSECRPRFRELVEFSRMAQDQRFVUJONEDLDGASPLDSTFYRSLLEDDMDKDLVDA	1020
0Y	1021	EEYLVPOQGFCCPDPAFGAGGMHHNRSSSTRSGGGDLTLGLPSEEBEAPRSPAPSEBG	1080
0Y	1021	EEYLVPOQGFCCPDPAFGAGGMHHNRSSSTRSGGGDLTLGLPSEEBEAPRSPAPSEBG	1080
0Y	745	AGSDVFGDGLMGAKGLOSLPTHDSPLQORYSEDPVLPBSETDGVYAPRLCSPQREYV	804

Db	1081	AGSDVFDSDLGKAGKAGQSDLPTHDPSLQRISEDPVPLPSETDGYAAPTCSQPEYV	1140
QY	805	NQPDVRPOPSPSREBPGLPAARPAAGATLLEPKTLSPGKNGVVDVAFGAVENPEYLPQ	864
Db	1141	NQPDVRPOPSPSREBPGLPAARPAAGATLLEAKTLSPGKNGVVDVAFGAVENPEYLPQ	1200
QY	865	GCALQPPHPPEAFSPAFDNLVYWDQDPPERCAPSPSTFKGTPTAENPEYLGIDPV	919
Db	1201	GCALQPPHPPEAFSPAFDNLVYWDQDPPERCAPSPSTFKGTPTAENPEYLGIDPV	1255
RESULT 6			
US-09-769-508-2			
Sequence 2, Application US/09769508			
Patent No. US20020155527A1			
GENERAL INFORMATION:			
APPLICANT: STUART, SUSAN G.			
APPLICANT: MONAHAN, JOHN J.			
APPLICANT: LANGTON, BEATRICE CLAUDIA			
APPLICANT: HANCOCK, MIRIAM E.C.			
APPLICANT: CHAO, LORRINE A.			
APPLICANT: BLUFORD, PETER			
TITLE OF INVENTION: C-ERBB-2 EXTERNAL DOMAIN: GP75			
FILE REFERENCE: BEBIO-111-C1			
CURRENT APPLICATION NUMBER: US/09/7769, 508			
CURRENT FILING DATE: 2001-01-26			
NUMBER OF SEQ ID NOS: 2			
SOFTWARE: PatentIn Ver. 2.1			
SEQ ID NO 2			
LENGTH: 1255			
TYPE: PRT			
ORGANISM: Homo sapiens			
US-09-769-508-2			
Query Match			
Best Local Similarity 95.4%, Score 4843; DB 9; Length 1255;			
Matches 913; Conservative 0; Mismatches 6; Indels 336; Gaps 1;			
QY	1	MELALACRWGILLALLPRGASASTQVCTGCTDMKRLPASPETHLMDLRHLYOGCQVVGNTL	60
Db	1	MELALACRWGILLALLPRGASASTQVCTGCTDMKRLPASPETHLMDLRHLYOGCQVVGNTL	60
QY	61	ELTYLPTNASTSLFDODIEVQGYVLIAHNOVQVYQLRRLRYRGQLFEDNVALAVLDNG	120
Db	61	ELTYLPTNASTSLFDODIEVQGYVLIAHNOVQVYQLRRLRYRGQLFEDNVALAVLDNG	120
QY	121	DPLNPTFVTVGASPGGLRELOLRSLTEILKGGVLLIQRPDOLCYODTILMKDIFHRKNOLA	180
Db	121	DPLNPTFVTVGASPGGLRELOLRSLTEILKGGVLLIQRPDOLCYODTILMKDIFHRKNOLA	180
QY	181	LTLIDTNSRACHPCSPCKCKSRGCGESSEDCQSILTRVYVACGACRCKGRLPTDCCHQOC	240
Db	181	LTLIDTNSRACHPCSPCKCKSRGCGESSEDCQSILTRVYVACGACRCKGRLPTDCCHQOC	240
QY	241	AAGCTGPRHSOCLACIHNHSGICELHCPALVYNTDFFESMPNDEGRYTFGASCVTACP	300
Db	241	AAGCTGPRHSOCLACIHNHSGICELHCPALVYNTDFFESMPNDEGRYTFGASCVTACP	300
QY	301	YNYLSTDVGSCCTVLCPLHNOEVTAEADGTORCEKCSKPCARVCYGLGMEHLREAVAVTSAN	360
Db	301	YNYLSTDVGSCCTVLCPLHNOEVTAEADGTORCEKCSKPCARVCYGLGMEHLREAVAVTSAN	360
QY	361	IOEFAGCKKIKGSLAFLEPESFDGDASTAPLQPPDOLVFEYLEBITGYLYISAMPDSLP	420
Db	361	IOEFAGCKKIKGSLAFLEPESFDGDASTAPLQPPDOLVFEYLEBITGYLYISAMPDSLP	420
QY	421	DLVSPONLOVIRGRILNHGAYSLTLOGGICISWLGRLSRELGGSLALLHNHTLCEFTVY	480
Db	421	DLVSPONLOVIRGRILNHGAYSLTLOGGICISWLGRLSRELGGSLALLHNHTLCEFTVY	480
QY	481	PWDOLFRRPHOALLHTANRPEDECVGEGSLACHQLCARGHCHWGPPTQCVCNCSQFLRGEC	540
Db	481	PWDOLFRRPHOALLHTANRPEDECVGEGSLACHQLCARGHCHWGPPTQCVCNCSQFLRGEC	540

Db 481 PWDOLFRRPHOALLHTANRPEDCEVGEGLACHOLCARALLGSGPTQCVCNSQFLRGDC 540  
QY 541 VEECRVLOGLPREYVNAHCLPCHEPCOPONGSVTCFGEPAOQVACAHYKDPPECVARC 600  
Db 541 VEECRVLOGLPREYVNAHCLPCHEPCOPONGSVTCFGEPAOQVACAHYKDPPECVARC 600  
QY 601 PSYGKPDLSYMPIMKFPDEBACOPCPINCHSCVDLDKCGPAQRASPLTS----- 653  
Db 601 PSYGKPDLSYMPIMKFPDEBACOPCPINCHSCVDLDKCGPAQRASPLTSIVSAVVG 660  
QY 654 ----- 653  
Db 661 ILLVVVLGVFGLIKRROQKIRKTYMRRLLOETELVEPLTPSGAMPNQAQMRILKETEL 720  
QY 654 ----- 653  
Db 721 RKKVLGSGAFYVKGIMIDGENVKIPVAIKVLENTSPKANKELDEAVVAGVSP 780  
QY 654 ----- 653  
Db 781 YVSRLLIGLSTVOLVTLQMPYGLLDHVRENRGLGSODLLMCMQIAKMSYLEDVR 840  
QY 654 ----- 653  
Db 841 LVHRDLAARNLVKSPNNVKITDEGLARLLIDETEHADGKVPDKMALESILRRFT 900  
QY 654 ----- 653  
Db 901 HSDVMSYGYVWELMTGAKPYDGIAPREIPDLLEKEGRLPQPICTIDVYMIWVKCM 960  
QY 654 ----- 684  
Db 961 IDSECRPRFRLVSEFSMARDPQRFVYIQNEDLGPASPLSTFYSRLLEDMDKDLVDA 1020  
QY 685 EEYLPOOGFCFDPAPAGGKVVHRRSSSTRSGGDLTLLESEBEAPRSLAPSEG 744  
Db 1021 EEYLPOOGFCFDPAPAGGKVVHRRSSSTRSGGDLTLLESEBEAPRSLAPSEG 1080  
QY 745 AGSDVFDDGLMGAAGKLSLPTHDSPLOKRSSEDPVPLPSETDGYVAPLTCSPQPEY 804  
Db 1081 AGSDVFDDGLMGAAGKLSLPTHDSPLOKRSSEDPVPLPSETDGYVAPLTCSPQPEY 1140  
QY 805 NOPDVRPQPPSPRSGPLPAARPAATLTERPKTLSPKNGVYKDVAFGGAENPEYLPQ 864  
Db 1141 NOPDVRPQPPSPRSGPLPAARPAATLTERPKTLSPKNGVYKDVAFGGAENPEYLPQ 1200  
QY 865 GGAAPQHPHPPAFSPAFDNLVYMODDPERGAPSTFKGTPTAENPEYLGIDVPY 919  
Db 1201 GGAAPQHPHPPAFSPAFDNLVYMODDPERGAPSTFKGTPTAENPEYLGIDVPY 1255

RESULT 7  
US-09-854-356-2  
Sequence 2, Application US/09854356  
Patent No. US20020177567A1  
GENERAL INFORMATION:  
APPLICANT: Cheever, Martin A.  
APPLICANT: Gheysen, Dirk  
APPLICANT: Corixa Corporation  
APPLICANT: SmithKline Beecham Biologicals S. A.  
TITLE OF INVENTION: HER-2/neu Fusion Proteins  
FILE REFERENCE: 014058-009810PC  
CURRENT APPLICATION NUMBER: US/09/854,356  
CURRENT FILING DATE: 2001-05-09  
PRIORITY APPLICATION NUMBER: US 09/493,480  
PRIORITY FILING DATE: 2000-01-28  
PRIOR APPLICATION NUMBER: US 60/117,976  
PRIOR FILING DATE: 1999-01-29  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 1256  
TYPE: PRT

ORGANISM: Rattus sp.  
FEATURE:  
OTHER INFORMATION: rat HER-2/neu protein  
NAME/KEY: DOMAIN  
LOCATION: (1)..(654)  
OTHER INFORMATION: extracellular domain (ECD)  
NAME/KEY: DOMAIN  
LOCATION: (677)..(1256)  
OTHER INFORMATION: intracellular domain (ICD)  
NAME/KEY: DOMAIN  
LOCATION: (721)..(998)  
OTHER INFORMATION: kinase domain (KD)  
NAME/KEY: DOMAIN  
LOCATION: (991)..(1256)  
OTHER INFORMATION: phosphorylation domain (PD)  
NAME/KEY: DOMAIN  
LOCATION: (991)..(1049)  
OTHER INFORMATION: fragment of the phosphorylation domain, preferred  
US-09-854-356-2

Query Match 81.5%; Score 4138.5; DB 9; Length 1256;  
Best Local Similarity 62.2%; Pred. No. 4,1e-231;  
Matches 781; Conservative 42; Mismatches 96; Indels 337; Gaps 2;

QY 1 MELALCRWGLLLALPPGAASTGYCTGTDKRLRPASPTHLDMLRHLVGGQVYQGNL 60  
Db 1 MELAMCRWGLLLALPPGIAAGTGYCTGTDKRLRPASPTHLDMLRHLVGGQVYQGNL 60  
QY 61 ELTYLPTNASLSFLDDIDYOGYVLIANNQVQVPLQRLRLRYRGQLFEDVYALAVLDNG 120  
Db 61 ELTYLPANASLSFLDDIDYOGYVLIANNQVQVPLQRLRLRYRGQLFEDVYALAVLDNR 120  
QY 121 DPLNNTTPYT-GASPGGLRELOLRSLTEILKGVLIQNNPOLCYODTILMKDIFHKNNQL 179  
Db 121 DPQDVAASTGRTRTEGLRELOLRSLTEILKGVLIQNNPOLCYODTILMKDIFHKNNQL 180  
QY 180 ALLTIDNRSRACHPSCPMKSGSRGWSSEDCOSLRTVACAGGACRKGPLPTDCHEQ 239  
Db 181 APVDIDNRSRACHPSCPMKSGSRGWSSEDCOSLRTVACAGGACRKGPLPTDCHEQ 240  
QY 240 CAAGCTGRKHSCLACLFHNSGTELCALPALTNTTFESMPRBEGRYTFGASCYVAC 299  
Db 241 CAAGCTGRKHSCLACLFHNSGTELCALPALTNTTFESMPRBEGRYTFGASCYVAC 300  
QY 300 PNYLSTVSGCTLYCPHNOEVTAEQNGOREKCSKFCARCYGLGMEHLREYAVATSA 359  
Db 301 PNYLSTVSGCTLYCPHNOEVTAEQNGOREKCSKFCARCYGLGMEHLREYAVATSD 360  
QY 360 NIOEPAGCKITFGSLAPLPESEFDGDPASNTAPLQEQLOVEFTEITGYLYISAMPDSL 419  
Db 361 NVQEFDCGCKITFGSLAPLPESEFDGDPASNTAPLQEQLOVEFTEITGYLYISAMPDSL 420  
QY 420 PDLSTFQNLQVYRGILNNGAYSLTLOGIGLSMLGLRSLRELGSGLALIHNTHTLCFYHT 479  
Db 421 RDLSTFQNLQVYRGILNNGAYSLTLOGIGLSMLGLRSLRELGSGLALIHNTHTLCFYHT 480  
QY 480 VPMDOLEFRNPQHALLHTANRPEDCEVGEGLACHOLCARHCGMGPPPTQVCNSQFLRGCE 539  
Db 481 VPMDOLEFRNPQHALLHTANRPEDCEVGEGLACHOLCARHCGMGPPPTQVCNSQFLRGCE 540  
QY 540 VEECRVLOGLPREYVNAHCLPCHEPCOPONGSVTCFGEPAOQVACAHYKDPPECVAR 599  
Db 541 VEECRVLOGLPREYVNAHCLPCHEPCOPONGSVTCFGEPAOQVACAHYKDPPECVAR 600  
QY 600 CPSGVKPDLSYMPIMKFPDEBACOPCPINCHSCVDLDKCGPAQRASPLT----- 652  
Db 601 CPSGVKPDLSYMPIMKFPDEBACOPCPINCHSCVDLDKCGPAQRASPLT----- 652  
QY 653 ----- 652  
Db 661 GVLLELLVVVVGIIIRKROKIRKTYMRRLLOETELVEPLTPSGAMPNQAQMRILKETE 720



QY 653 ----- 652  
Db 721 LKRVKVLGSGAGTGYKGIWIPDGENVKIIPVAKILRENTSPKANKEILDEAYVAGVG 780  
QY 653 ----- 652  
Db 781 PYVSRLGICLSTVOLVQMLPYGCLLDHVRHGRGLSODLLNMCVQIAKMSYLEDY 840  
QY 653 ----- 652  
Db 841 RLVRHDLAARNVLKSPNHVKITDFGLARLDIDETEHADGKVPDKMMALESLRRRF 900  
QY 653 ----- 652  
Db 901 THOSDWSYGVTVWELMTFGAKPYDGIIPARELPDLEKEGERLPORPCTIDVYIMVYKCM 960  
QY 653 ----- 683  
Db 961 MIDSECRPRFRELVSFESHARDPQRFVYIQNEDLGSSPMSTFVRSLLDEDDMGDLVD 1020  
QY 684 ABEYLVPOQGFPCPPAPAGAGVHHRRSSSTRSGGDLTLGLEPSEEPAPRSLAPSE 743  
Db 1021 ABEYLVPOQGFSPDPPTGSTAHRRSSSTRSGGDLTLGLEPSEEGPRSPAPSE 1080  
QY 744 GAGSDVFDGDLGAKGLOSLPTHDPSPLOHRYSEDPTVPLPSEHTDGYVAPLTCSPOPEY 803  
Db 1081 GAGSDVFDGDLGAKGLOSLPTHDPSPLOHRYSEDPTVPLPSEHTDGYVAPLTCSPOPEY 1140  
QY 804 VNQPDVROPSPREGPLCPAARPAAGATLERPKTLSPKNGVYKDVAFAGAVENPEYLP 863  
Db 1141 VNQSEVQOPPLTPREGPLCPAARPAAGATLERPKTLSPKNGVYKDVAFAGAVENPEYLP 1200  
QY 864 OGGAPOHPHPAPSPAFDNLTYWDODPREPGARPSTFGTSPANPEYLGDLVY 919  
Db 1201 REGTASPPHPAPSPAFDNLTYWDONSSSEOGPPSPNEGTPTAENPEYLGDLVY 1256

## RESULT 8

US-09-870-759-118

Sequence 118, Application US/09870759

Patent No. US20020177551A1

GENERAL INFORMATION:

APPLICANT: TERMAN, David S

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE

FILE REFERENCE: 870759

CURRENT APPLICATION NUMBER: US/09/870,759

PRIOR FILING DATE: 2002-01-14

PRIOR APPLICATION NUMBER: US 60/208,128

NUMBER OF SEQ ID NOS: 166

SOFTWARE: PatentIn version 3.1

SEQ ID NO 118

LENGTH: 1260

TYPE: PRT

ORGANISM: Rattus norvegicus

US-09-870-759-118

Query Match 81.4%; Score 4134; DB 9; Length 1260;  
Best Local Similarity 62.1%; Pred. No. 7.4e-231;  
Matches 781; Conservative 42; Mismatches 96; Indels 338; Gaps 3;

QY 1 METAAALCRGGLLALLPAGASTOVCTGDMKRLPASPETHLDMRHLTYOGCOVYQGM 60  
Db 4 MELAAACRMGFLIALLPPIAGTGYCTGDMKRLPASPETHLDMRHLTYOGCOVYQGM 63  
QY 61 ELTYLPTNLSLFLDIOEVQGVYLIANQVROVPLQRLIRVIGTQLDFEDNTALAVLDNG 120  
Db 64 ELTYVPANASLFLDIOEVQGVYLIANQVRYVPLQRLIRVIGTQLDFEDKALAVLDNR 123  
QY 121 DPLANTTPVT-GASPGGLRELDRLSTELKAGVLIQNRPOLCYODTLIMKDIFFKNNOL 179  
Db 124 DPQDNVAASTPGRTPEGLRELDRLSTELKAGVLIQNRPOLCYODMVLIMKDVFFKNNOL 183

QY 180 ALTLIDNRSRACHPCSPMCKSGRCWGESSEDDQSLTFTVACAGGACRCKGPLPTDCCHEO 239  
Db 184 APYDIDNRSRACPCAPACKDHCWGESSEDDQSLTFTVACAGGACRCKGPLPTDCCHEO 243  
QY 240 CAAGCTGPKHSDIACLHNHSGICELHC PALVYNTDFEESMNPREGRTYGASCVTAC 299  
Db 244 CAAGCTGPKHSDIACLHNHSGICELHC PALVYNTDFEESMNPREGRTYGASCVTAC 303  
QY 300 PNYVLTSDVSGCTLVCPRLNQEYTAEDGTQCEKSCPCARVCYGLGMEHLREAVATSA 359  
Db 304 PNYVLTSEVSGCTLVCPRLNQEYTAEDGTQCEKSCPCARVCYGLGMEHLREAVATSA 363  
QY 360 NIOEFAGCKKITGSLAFLEPESFDGDPASNTAPLQPOLOVFTLEITITGYLISAPDLSL 419  
Db 364 NIOEFAGCKKITGSLAFLEPESFDGDPASNTAPLQPOLOVFTLEITITGYLISAPDLSL 423  
QY 420 PDLVSQNLQVIRGRILHNGAVSLTLQGGIISWLGIRSLRELGSGALITHNHTLCEVHT 479  
Db 424 RDLVSQNLQVIRGRILHNGAVSLTLQGGIISWLGIRSLRELGSGALITHNHTLCEVHT 483  
QY 480 VPMDQLFRNPQALHTANRPEDE-CVGEGLACHOLICANGHCWGPPTQCVNCSOFLRQ 538  
Db 484 VPMDQLFRNPQALHTANRPEDE-CVGEGLACHOLICANGHCWGPPTQCVNCSOFLRQ 543  
QY 539 ECVSECRVLOGLPREVYNARHCLPCHPECOPONGSVTCGPAPADOCVAAHKKDPEFCVA 598  
Db 544 ECVSECRVLOGLPREVYNARHCLPCHPECOPONGSVTCGPAPADOCVAAHKKDPEFCVA 603  
QY 599 RCPSGVCPDLSTYMPIMKFPDEGACPCP INCTHSCVDLDDKCPAEGNASPLT----- 652  
Db 604 RCPSGVCPDLSTYMPIMKFPDEGACPCP INCTHSCVDLDDKCPAEGNASPLT----- 663  
QY 653 ----- 652  
Db 664 EGVLLFLLVVYVGLIKRRROKIRKTYMRLLQETELVEPLTPSGAMPNOAMRLKET 723  
QY 653 ----- 652  
Db 724 ELRKVYVLGSGAGTGYKGIWIPDGENVKIIPVAKILRENTSPKANKEILDEAYVAGVG 783  
QY 653 ----- 652  
Db 784 SPVSRLLGICLSTVOLVQMLPYGCLLDHVRHGRGLSODLLNMCVQIAKMSYLEDY 843  
QY 653 ----- 652  
Db 844 VRLVRHDLAARNVLKSPNHVKITDFGLARLDIDETEHADGKVPDKMMALESLRRRF 903  
QY 653 ----- 652  
Db 904 FTHOSDWSYGVTVWELMTFGAKPYDGIIPARELPDLEKEGERLPORPCTIDVYIMVYKCM 963  
QY 653 ----- 682  
Db 964 MIDSECRPRFRELVSFESHARDPQRFVYIQNEDLGSSPMSTFVRSLLDEDDMGDLVD 1023  
QY 683 DAEBEYLVPOQGFPCPPAPAGAGVHHRRSSSTRSGGDLTLGLEPSEEPAPRSLAPSE 742  
Db 1024 DAEBEYLVPOQGFSPDPPTGSTAHRRSSSTRSGGDLTLGLEPSEEGPRSPAPSE 1083  
QY 743 EGAGSDVFDGDLGAKGLOSLPTHDPSPLOHRYSEDPTVPLPSEHTDGYVAPLTCSPOPE 802  
Db 1084 EGAGSDVFDGDLGAKGLOSLPTHDPSPLOHRYSEDPTVPLPSEHTDGYVAPLTCSPOPE 1143  
QY 803 YVNPDPVROPSPREGPLCPAARPAAGATLERPKTLSPKNGVYKDVAFAGAVENPEYLP 862  
Db 1144 YVNPDPVROPSPREGPLCPAARPAAGATLERPKTLSPKNGVYKDVAFAGAVENPEYLP 1203  
QY 863 POGGAPOHPHPAPSPAFDNLTYWDODPREPGARPSTFGTSPANPEYLGDLVY 919  
Db 1204 PREGTASPPHPAPSPAFDNLTYWDONSSSEOGPPSPNEGTPTAENPEYLGDLVY 1260

RESULT 9  
US-09-854-356-14  
; Sequence 14, Application US/09854356  
; Patent No. US2002017567A1  
; GENERAL INFORMATION:  
; APPLICANT: Cheever, Martin A.  
; APPLICANT: Gheysen, Dirk  
; APPLICANT: Corixa Corporation  
; APPLICANT: SmithKline Beecham Biologicals S. A.  
; TITLE OF INVENTION: HER-2/neu Fusion Proteins  
; FILE REFERENCE: 014058-009810PC  
; CURRENT APPLICATION NUMBER: US/09/854,356  
; CURRENT FILING DATE: 2001-05-09  
; PRIOR APPLICATION NUMBER: US 09/493,480  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: US 60/117,976  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: Patentl Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 1256  
; TYPE: PRT  
; ORGANISM: Mus sp.  
; FEATURE:  
; HER INFORMATION: mouse HER-2/neu protein  
US 854-356-14  
; QTY Match 81.2%; Score 4125.5; DB 9; Length 1256;  
; Bit Local Similarity 62.1%; Pred. No. 2.3e-230;  
; Matches 780; Conservative 44; Mismatches 95; Indels 337; Gaps 2;  
QY 1 MELALCRWGLLALLPFGASTQVCTGTDMLRLPASPETHDMLRLHYGCGVOVGNL 60  
DB 1 MELAMCRWGLLALLPFGASTQVCTGTDMLRLPASPETHDMLRLHYGCGVOVGNL 60  
QY 61 ELTLPTASLSFLQDIOEVGYVLIANOVROYPLQRLRYKGTOLFEDNYALAVLDNG 120  
DB 61 ELTLPTASLSFLQDIOEVGYVLIANOVROYPLQRLRYKGTOLFEDNYALAVLDNR 120  
QY 121 DPLNN-TPPVYAGPGGLREIQLRSLEIILKGVLIQNPOLCYODITLMDIFHKNNQL 179  
DB 121 DPLNN-TPPVYAGPGGLREIQLRSLEIILKGVLIQNPOLCYODITLMDIFHKNNQL 179  
QY 121 DPLNN-TPPVYAGPGGLREIQLRSLEIILKGVLIQNPOLCYODITLMDIFHKNNQL 180  
DB 121 DPLNN-TPPVYAGPGGLREIQLRSLEIILKGVLIQNPOLCYODITLMDIFHKNNQL 180  
QY 180 ALLLIDNRSRACHPCSPMCKSGRCWSESSDCSLARTVAGGACARCKGRLPTDCCHEQ 239  
DB 180 ALLLIDNRSRACHPCSPMCKSGRCWSESSDCSLARTVAGGACARCKGRLPTDCCHEQ 239  
QY 181 APVMDTRSRACPPCAPTCKDNHCWSESPEDCOLITGCTSGCARCKGRPLDCCHEQ 240  
DB 181 APVMDTRSRACPPCAPTCKDNHCWSESPEDCOLITGCTSGCARCKGRPLDCCHEQ 240  
QY 240 CAAGCTGKHSDDLACLFHNSGICELHCPALVTYNTDTFESMNPBGRYTFGASCVTATC 299  
DB 240 CAAGCTGKHSDDLACLFHNSGICELHCPALVTYNTDTFESMNPBGRYTFGASCVTATC 299  
QY 300 PYNLSTDVSGCTVCPPLHNOEVAEDTQRCERCKSPCARVCYGLGMEHLREYRAVTS 359  
DB 300 PYNLSTDVSGCTVCPPLHNOEVAEDTQRCERCKSPCARVCYGLGMEHLREYRAVTS 359  
QY 301 PYNLSTEVSGCTVCPPLHNOEVAEDTQRCERCKSPCARVCYGLGMEHLREYRAVTS 360  
DB 301 PYNLSTEVSGCTVCPPLHNOEVAEDTQRCERCKSPCARVCYGLGMEHLREYRAVTS 360  
QY 360 NIOFAGCKTFGSLAPESFDGDPASNTAPLOPEOLQVEFTELETTGYLYISAMPDSL 419  
DB 360 NIOFAGCKTFGSLAPESFDGDPASNTAPLOPEOLQVEFTELETTGYLYISAMPDSL 419  
QY 361 NIOFAGCKTFGSLAPESFDGDPASNTAPLOPEOLQVEFTELETTGYLYISAMPDSL 420  
DB 361 NIOFAGCKTFGSLAPESFDGDPASNTAPLOPEOLQVEFTELETTGYLYISAMPDSL 420  
QY 420 PDLVFNLOVYIRGRILLHNGVYSLTLOGISWLGIRSLRELSGLALIHNTHLCPVHT 479  
DB 420 PDLVFNLOVYIRGRILLHNGVYSLTLOGISWLGIRSLRELSGLALIHNTHLCPVHT 479  
QY 421 QDLVFNLOVYIRGRILLHNGVYSLTLOGISWLGIRSLRELSGLALIHNTHLCPVHT 480  
DB 421 QDLVFNLOVYIRGRILLHNGVYSLTLOGISWLGIRSLRELSGLALIHNTHLCPVHT 480  
QY 480 VPMQDLFRNPQALHTANRPEDCVGEGLAGHOLCAARGHGMGPGTQCVNCSOFLRGOE 539  
DB 480 VPMQDLFRNPQALHTANRPEDCVGEGLAGHOLCAARGHGMGPGTQCVNCSOFLRGOE 539  
QY 481 VPMQDLFRNPQALHTANRPEDCVGEGLAGHOLCAARGHGMGPGTQCVNCSOFLRGOE 540  
DB 481 VPMQDLFRNPQALHTANRPEDCVGEGLAGHOLCAARGHGMGPGTQCVNCSOFLRGOE 540  
QY 540 CVEBCRVQLGRLPREYVNAHCLPCHPECOPONGSVTCFGPAPDOCVACAHYKDPFCVAR 599  
DB 540 CVEBCRVQLGRLPREYVNAHCLPCHPECOPONGSVTCFGPAPDOCVACAHYKDPFCVAR 599  
QY 541 CVEBCRVQLGRLPREYVNAHCLPCHPECOPONGSVTCFGPAPDOCVACAHYKDPFCVAR 600  
DB 541 CVEBCRVQLGRLPREYVNAHCLPCHPECOPONGSVTCFGPAPDOCVACAHYKDPFCVAR 600  
QY 600 CPSSGVKPDLSVMPIMKPFDEGACOPCPINCTHSCVLDLDDKCPAPDOCVACAHYKDPFCVAR 652  
DB 600 CPSSGVKPDLSVMPIMKPFDEGACOPCPINCTHSCVLDLDDKCPAPDOCVACAHYKDPFCVAR 652

DB 601 CPSSGVKPDLSVMPIMKPFDEGACOPCPINCTHSCVLDLDDKCPAPDOCVACAHYKDPFCVAR 660  
QY 653 -----  
DB 661 GVLLFLIIIVVIGILIKRRROKIKKYMRLQETELVEPLTPSGAVPNOAQMILKETE 720  
QY 653 -----  
DB 721 LRLKLVLGSAFGTVYKGIWIPDGENYKIPYAIKVLRENTSPKANKELDEAYVAVGVS 780  
QY 653 -----  
DB 781 PYVSLGLICLTSTVOLYQMLPYGCLLDHYREHRLGSDLLMKVCQJAKGMSTLEEY 840  
QY 653 -----  
DB 841 RLVRDLAARVLYKSPNHVYITPFGARLLDIDETEVHADGKVPYIKWMALESILRRF 900  
QY 653 -----  
DB 901 THQSDVMSYGVTVMLMTFGAKPYDGIPAREIPDLLEKGERLPQPICTIDVYMINVKCW 960  
QY 653 -----  
DB 961 MIDSECRPRRELVSERSMARDPQRFVYQNEDELGSSPDSTFYRSLLLEDDMGELVD 1020  
QY 684 AEELVLPQOGFPCPDPAAGAGVYHNRSSSTRSGGDLTLGLPESEEARSPPLAPSE 743  
DB 1021 AEELVLPQOGFPCPDPAAGAGVYHNRSSSTRSGGDLTLGLPESEEARSPPLAPSE 1080  
QY 744 GAGSDVEFDGIDGKAARGLSLPTHPSPLOYSDEPTVLPSTDTGYVAPLTCSPQPEY 803  
DB 1081 GAGSDVEFDGIDGKAARGLSLPTHPSPLOYSDEPTVLPSTDTGYVAPLTCSPQPEY 1140  
QY 804 VNQPDVPRQPSRREGPLPAARPGATLERKTLSPKNGVYKVFAPGAVENPEYLP 863  
DB 1141 VNQPDVPRQPSRREGPLPAARPGATLERKTLSPKNGVYKVFAPGAVENPEYLP 1200  
QY 864 QGGAAPQHPHPAPAFNDLYYWDOPPERGAPSPFTFGTAENPEYGLDVPV 919  
DB 1201 QGGAAPQHPHPAPAFNDLYYWDOPPERGAPSPFTFGTAENPEYGLDVPV 1256  
RESULT 10  
US-09-854-356-7  
; Sequence 7, Application US/09854356  
; Patent No. US2002017567A1  
; GENERAL INFORMATION:  
; APPLICANT: Cheever, Martin A.  
; APPLICANT: Gheysen, Dirk  
; APPLICANT: Corixa Corporation  
; APPLICANT: SmithKline Beecham Biologicals S. A.  
; TITLE OF INVENTION: HER-2/neu Fusion Proteins  
; FILE REFERENCE: 014058-009810PC  
; CURRENT APPLICATION NUMBER: US/09/854,356  
; CURRENT FILING DATE: 2001-05-09  
; PRIOR APPLICATION NUMBER: US 09/493,480  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: US 60/117,976  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: Patentl Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 712  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: fusion protein  
; OTHER INFORMATION: of BCD and delta PD of human HER-2/neu  
US-09-854-356-7  
; QTY Match 77.9%; Score 3954; DB 9; Length 712;  
; Best Local Similarity 100.0%; Pred. No. 9.2e-221;

	Matches	712:	Conservative	0:	Mismatches	0:	Indels	0:	Gaps	0:
QY	1	MELAA	CSWGLLALLPRGA	ASTQVCTGTDMKRLRPAS	ETHLDMRLHYG	CCQVVGNTL	60			
Db	1	MELAA	CSWGLLALLPRGA	ASTQVCTGTDMKRLRPAS	ETHLDMRLHYG	CCQVVGNTL	60			
QY	61	ELTYPT	NASISLFODIO	ENGUYVLIANOV	ROYLQRLRYRG	QLTEEDN	ALAVLDNG	120		
Db	61	ELTYPT	NASISLFODIO	ENGUYVLIANOV	ROYLQRLRYRG	QLTEEDN	ALAVLDNG	120		
QY	121	DPLNNT	TVTVTASPGGL	RELDRLSR	TELLKGVLLQNR	PNOLCYOQTILMKD	LFHKNNOA	180		
Db	121	DPLNNT	TVTVTASPGGL	RELDRLSR	TELLKGVLLQNR	PNOLCYOQTILMKD	LFHKNNOA	180		
QY	181	LTLIDT	NRSRACHPCSP	CKSCRCWGESSE	EDQSLTFTVCA	GAGCARCKGRLPTD	CCHEQC	240		
Db	181	LTLIDT	NRSRACHPCSP	CKSCRCWGESSE	EDQSLTFTVCA	GAGCARCKGRLPTD	CCHEQC	240		
QY	241	AAGCTG	PRHSOCLAH	NHSGICLHCP	ALVYTYTDTTFE	SPNREGRYTVG	ASCVTACP	300		
Db	241	AAGCTG	PRHSOCLAH	NHSGICLHCP	ALVYTYTDTTFE	SPNREGRYTVG	ASCVTACP	300		
QY	301	YNYLST	DVSGCTLVCP	LNOEVTADG	TQRECKSKPCAR	VCYGLGMEHLR	VRVAVTSAN	360		
Db	301	YNYLST	DVSGCTLVCP	LNOEVTADG	TQRECKSKPCAR	VCYGLGMEHLR	VRVAVTSAN	360		
QY	361	IOEFAG	CKKIRGSLAF	PESEFDGP	ASNTARLPOLQ	VFETLEETIGYLYI	SAMPDILP	420		
Db	361	IOEFAG	CKKIRGSLAF	PESEFDGP	ASNTARLPOLQ	VFETLEETIGYLYI	SAMPDILP	420		
QY	421	DLSVFQ	NQVIRGRIL	NGAYSLT	LOGIGISWLG	RSIRELGSGIAL	HNHTHLCFVTV	480		
Db	421	DLSVFQ	NQVIRGRIL	NGAYSLT	LOGIGISWLG	RSIRELGSGIAL	HNHTHLCFVTV	480		
QY	481	PMDOLE	RNP	PHALLT	TANRPEDEC	GESEGLACHOLCAR	HCMPGPTQSCVNC	SQFLRGQEC	540	
Db	481	PMDOLE	RNP	PHALLT	TANRPEDEC	GESEGLACHOLCAR	HCMPGPTQSCVNC	SQFLRGQEC	540	
QY	541	VEECGR	VLOGLPRE	YVNAHNC	LCPRHECQ	PONGSVTCFGR	PADQCYA	CAHXYKDRP	CVARC	600
Db	541	VEECGR	VLOGLPRE	YVNAHNC	LCPRHECQ	PONGSVTCFGR	PADQCYA	CAHXYKDRP	CVARC	600
QY	601	PSGYK	PDLSTYR	PIKFPDE	BGACOPCP	INCTHSCYD	LDKGRABQ	RSAPLTSQ	NEDLGP	660
Db	601	PSGYK	PDLSTYR	PIKFPDE	BGACOPCP	INCTHSCYD	LDKGRABQ	RSAPLTSQ	NEDLGP	660
QY	661	ASPLDST	YRSLL	EDDDKGL	VDAAE	YLVPOOG	FCRPRAC	GAGM	VNHNRR	712
Db	661	ASPLDST	YRSLL	EDDDKGL	VDAAE	YLVPOOG	FCRPRAC	GAGM	VNHNRR	712
RESULT 11										
US-09-854-356-3										
Sequence 3, Application US/09854356										
Patent No. US2002017567A1										
GENERAL INFORMATION:										
APPLICANT: Cheever, Martin A.										
APPLICANT: Cheever, Martin A.										
APPLICANT: Corixa Corporation										
TITLE OF INVENTION: HER-2/neu Fusion Proteins										
FILE REFERENCE: 014058-009810PC										
CURRENT APPLICATION NUMBER: US/09/854,356										
CURRENT FILING DATE: 2001-05-09										
PRIOR APPLICATION NUMBER: US 09/493,480										
PRIOR FILING DATE: 2000-01-28										
PRIOR APPLICATION NUMBER: US 60/117,976										

```

; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: extracellular domain (ECD) of human HER-2/neu
US-09-854-356-3

Query Match          71.4%: Score 3628; DB 9; Length 653;
Best Local Similarity 100.0%: Pred. No. 5,1e-202;
Matches 653: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MEIALLCRNGLLIALLPFGAASSTGYCTGDMKLRYPASPETHIDMLRHLGYGCGVVGNTL 60
Db 1 MEIALLCRNGLLIALLPFGAASSTGYCTGDMKLRYPASPETHIDMLRHLGYGCGVVGNTL 60

61 ELTYLPFNASLSFLDDIQEVGVYVLIANNQVQVPLQRLRLIRIVGTQLFEDNYALAVLDNG 120
Db 61 ELTYLPFNASLSFLDDIQEVGVYVLIANNQVQVPLQRLRLIRIVGTQLFEDNYALAVLDNG 120

121 DPLNNTPTPTGASPGGLRRLQLRSLEILKGGVLIQRRNQQLCYOTIILMKDIFHNKNNOLA 180
Db 121 DPLNNTPTPTGASPGGLRRLQLRSLEILKGGVLIQRRNQQLCYOTIILMKDIFHNKNNOLA 180

181 LFLIDPNRSRAACHPCSPMKGSRGCGESSEDDQSILRTYACAGACARCKGRLPTDCHHEQC 240
Db 181 LFLIDPNRSRAACHPCSPMKGSRGCGESSEDDQSILRTYACAGACARCKGRLPTDCHHEQC 240

241 AAGCTGPKHSDDLACILHFNHSGICELHCPALVTYNTDFESMPNDEGRYTFGASCVTACP 300
Db 241 AAGCTGPKHSDDLACILHFNHSGICELHCPALVTYNTDFESMPNDEGRYTFGASCVTACP 300

301 YVYLSIDVSSCLVCPRLHNOEYTAEDGYQRCRCKSKPCARVCTGIGEMELREYRAVTSAN 360
Db 301 YVYLSIDVSSCLVCPRLHNOEYTAEDGYQRCRCKSKPCARVCTGIGEMELREYRAVTSAN 360

361 IOEFAGCKRTFCSLAFLPRESFGDDPASNTAPQPPOLQVFEETLEETGGLYTSAMPDSL 420
Db 361 IOEFAGCKRTFCSLAFLPRESFGDDPASNTAPQPPOLQVFEETLEETGGLYTSAMPDSL 420

421 DLSVFQNLQVIRGRILHNGAYSLLDQIGISWLGRLSRELSGSLALIHNTHLQCFVHTV 480
Db 421 DLSVFQNLQVIRGRILHNGAYSLLDQIGISWLGRLSRELSGSLALIHNTHLQCFVHTV 480

481 PWDOLFRRNHQALLHTANRPDEECYGEGLACIQLCARHGCMRGPTQVCNCSQFLRGDEC 540
Db 481 PWDOLFRRNHQALLHTANRPDEECYGEGLACIQLCARHGCMRGPTQVCNCSQFLRGDEC 540

541 VEECRVLQGLPREYVNAHRCILCPHRCOPQNSVTCFGEADQCYAACAHYKPRPCVARC 600
Db 541 VEECRVLQGLPREYVNAHRCILCPHRCOPQNSVTCFGEADQCYAACAHYKPRPCVARC 600

601 PGGVPRDLSYMIWKFPDEGACQCPINCTHSCVDLDDKGPAPQRRASPLTS 653
Db 601 PGGVPRDLSYMIWKFPDEGACQCPINCTHSCVDLDDKGPAPQRRASPLTS 653

RESULT 12
US-09-921-161-1
; Sequence 1, Application US/09921161
; Patent No. US2002090662A1
; GENERAL INFORMATION:
; APPLICANT: Ralph, Peter
; TITLE OF INVENTION: ANALYTICAL METHOD
; FILE REFERENCE: GENE.066A
; CURRENT APPLICATION NUMBER: US/09/921,161
; CURRENT FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/225,433
; PRIOR FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 645
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-921-161-1

```

Query Match 70.7%; Score 3590; DB 10; Length 645;  
Best Local Similarity 100.0%; Pred. No. 7.8e-200;  
Matches 645; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MELAAACRMGGLIALPPGAASVCTGTGDMKRLPASPETHLMDRLHLYGGCCVYQGNL 60
DB 1 MELAAACRMGGLIALPPGAASVCTGTGDMKRLPASPETHLMDRLHLYGGCCVYQGNL 60
QY 61 ELTYLPTNASSFLQDIOEVGYLIAHNOYRQVPLQRLRIYRGTLQFEDNVALAVLNG 120
DB 61 ELTYLPTNASSFLQDIOEVGYLIAHNOYRQVPLQRLRIYRGTLQFEDNVALAVLNG 120
QY 121 DPLNTPVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYODTILMKDIFHKNNOL 180
DB 121 DPLNTPVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYODTILMKDIFHKNNOL 180
QY 181 LTLIDTNRSRACHPCSPMKSGRCWGESSEDCQSLTRVYAGGACRCKGRLPTDCCHEQ 240
DB 181 LTLIDTNRSRACHPCSPMKSGRCWGESSEDCQSLTRVYAGGACRCKGRLPTDCCHEQ 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNDEGRYTFGASCVTAC 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNDEGRYTFGASCVTAC 300
QY 301 YNYLSTDVGSCTLYCPLHNOEYTAEDGTQRCCKSKPCARVCYGLMEHLREVAVTSA 360
DB 301 YNYLSTDVGSCTLYCPLHNOEYTAEDGTQRCCKSKPCARVCYGLMEHLREVAVTSA 360
QY 361 IOEFAGCKKIFGSLAFLEPSEFDGPASNTAPLOPEQLOVEFTELEITGYLYISAMPDSL 420
DB 361 IOEFAGCKKIFGSLAFLEPSEFDGPASNTAPLOPEQLOVEFTELEITGYLYISAMPDSL 420
QY 421 DLSYFQNLQVIRGRILHNGAYSLTLOGLISWLGRLSLRELGSGLALIHNTLCEVHTY 480
DB 421 DLSYFQNLQVIRGRILHNGAYSLTLOGLISWLGRLSLRELGSGLALIHNTLCEVHTY 480
QY 481 PWDOLFRRNPQALHTANRPDECEVGEGLACHQCLCARGHGCGPPTQCVNCSQFLRGQEC 540
DB 481 PWDOLFRRNPQALHTANRPDECEVGEGLACHQCLCARGHGCGPPTQCVNCSQFLRGQEC 540
QY 541 VEECVLQGLPREYVNAHRCPLCHPECOPONGSVTCFGEADQCAACAHYKDPPECVAV 600
DB 541 VEECVLQGLPREYVNAHRCPLCHPECOPONGSVTCFGEADQCAACAHYKDPPECVAV 600
QY 601 PSQVPRDLSTYPIWKFPDEEGACQPCPINCTHSCVDLDDKGCAPAE 645
DB 601 PSQVPRDLSTYPIWKFPDEEGACQPCPINCTHSCVDLDDKGCAPAE 645
```

RESULT 13  
US-09-854-356-8  
Sequence 8, Application US/09854356  
Patent No. US2002017567A1

GENERAL INFORMATION:  
APPLICANT: Cheever, Martin A.  
APPLICANT: Gheysen, Dirk  
APPLICANT: Corixa Corporation  
APPLICANT: SmithKline Beecham Biologicals S. A.  
TITLE OF INVENTION: HER-2/neu Fusion Proteins  
FILE REFERENCE: 014058-009810PC  
CURRENT APPLICATION NUMBER: US/09/854,356  
CURRENT FILING DATE: 2001-05-09  
PRIORITY APPLICATION NUMBER: US 09/493,480  
PRIOR FILING DATE: 2000-01-28  
PRIOR APPLICATION NUMBER: US 60/117,976  
PRIOR FILING DATE: 1999-01-29  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 8  
LENGTH: 654  
TYPE: PRT  
ORGANISM: Rattus sp.

FEATURE:  
OTHER INFORMATION: extracellular domain (ECD) of rat HER-2/neu  
US-09-854-356-8

Query Match 61.3%; Score 3110.5; DB 9; Length 654;  
Best Local Similarity 85.5%; Pred. No. 3.4e-112;  
Matches 558; Conservative 32; Mismatches 62; Indels 1; Gaps 1;

```
QY 1 MELAAACRMGGLIALPPGAASVCTGTGDMKRLPASPETHLMDRLHLYGGCCVYQGNL 60
DB 1 MELAAACRMGGLIALPPGAASVCTGTGDMKRLPASPETHLMDRLHLYGGCCVYQGNL 60
QY 61 ELTYLPTNASSFLQDIOEVGYLIAHNOYRQVPLQRLRIYRGTLQFEDNVALAVLNG 120
DB 61 ELTYLPTNASSFLQDIOEVGYLIAHNOYRQVPLQRLRIYRGTLQFEDNVALAVLNG 120
QY 121 DPLNTPVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYODTILMKDIFHKNNOL 179
DB 121 DPLNTPVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYODTILMKDIFHKNNOL 180
QY 180 ALTLIDTNRSRACHPCSPMKSGRCWGESSEDCQSLTRVYAGGACRCKGRLPTDCCHEQ 239
DB 180 ALTLIDTNRSRACHPCSPMKSGRCWGESSEDCQSLTRVYAGGACRCKGRLPTDCCHEQ 240
QY 240 CAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNDEGRYTFGASCVTAC 299
DB 240 CAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNDEGRYTFGASCVTAC 300
QY 300 PNYLSTDVGSCTLYCPLHNOEYTAEDGTQRCCKSKPCARVCYGLMEHLREVAVTSA 359
DB 300 PNYLSTDVGSCTLYCPLHNOEYTAEDGTQRCCKSKPCARVCYGLMEHLREVAVTSA 360
QY 360 NIOEFAGCKKIFGSLAFLEPSEFDGPASNTAPLOPEQLOVEFTELEITGYLYISAMPDSL 419
DB 360 NIOEFAGCKKIFGSLAFLEPSEFDGPSSGIAPLREQLQVEFTELEITGYLYISAMPDSL 420
QY 420 PDLSTYFQNLQVIRGRILHNGAYSLTLOGLISWLGRLSLRELGSGLALIHNTLCEVHTY 479
DB 420 PDLSTYFQNLQVIRGRILHNGAYSLTLOGLISWLGRLSLRELGSGLALIHNTLCEVHTY 480
QY 480 VPDOLFRRNPQALHTANRPDECEVGEGLACHQCLCARGHGCGPPTQCVNCSQFLRGQEC 539
DB 480 VPDOLFRRNPQALHTANRPDECEVGEGLACHQCLCARGHGCGPPTQCVNCSQFLRGQEC 540
QY 540 CVEECRVLQGLPREYVNAHRCPLCHPECOPONGSVTCFGEADQCAACAHYKDPPECVAV 599
DB 540 CVEECRVLQGLPREYVNAHRCPLCHPECOPONGSVTCFGEADQCAACAHYKDPPECVAV 600
QY 600 CPSGVPRDLSTYPIWKFPDEEGACQPCPINCTHSCVDLDDKGCAPAEORASPT 652
DB 600 CPSGVPRDLSTYPIWKFPDEEGACQPCPINCTHSCVDLDDKGCAPAEORASPT 653
```

RESULT 14

US-09-821-883-3  
Sequence 3, Application US/09821883  
Patent No. US20020061310A1

GENERAL INFORMATION:  
APPLICANT: Vidovic, Damir  
APPLICANT: Graddis, Thomas  
TITLE OF INVENTION: Compositions and Methods for Dendritic  
FILE REFERENCE: 7636-0022.30  
CURRENT APPLICATION NUMBER: US/09/821,883  
CURRENT FILING DATE: 2001-03-30  
PRIOR APPLICATION NUMBER: US 60/193,504  
PRIOR FILING DATE: 2000-03-30  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 564  
TYPE: PRT

ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: HER500\* construct  
 US-09-821-883-3

Query Match 50.9%; Score 2585; DB 10; Length 564;  
 Best Local Similarity 56.7%; Pred. No. 5,6e-142;  
 Matches 512; Conservative 3; Mismatches 6; Indels 382; Gaps 2;

16 LPPGASTVCTGTDKMLRLPASPEHMLRLHLYGGCOVVOGNLELTLPNTASISFQ 75  
 35 LARGAASVCTGTDKMLRLPASPEHMLRLHLYGGCOVVOGNLELTLPNTASISFQ 94  
 76 DIOEVGYVLIANOVROYPLQRLRIYRGTOLEFEDNYALAVLDNGPPLNNTPVYTGASPG 135  
 95 DIOEVGYVLIANOVROYPLQRLRIYRGTOLEFEDNYALAVLDNGPPLNNTPVYTGASPG 154  
 136 GLREIOLRLSTELIKGVLIOIRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPC 195  
 155 GLREIOLRLSTELIKGVLIOIRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPC 214  
 196 SPWCKSGRWGSSSEDCOSLTRVCAGGACRCKGRLPTDCHEQCAAGCTGPKHSCLAC 255  
 215 SPWCKSGRWGSSSEDCOSLTRVCAGGACRCKGRLPTDCHEQCAAGCTGPKHSCLAC 274  
 256 LHPNHSIGICELCPALVTYNTDFFESMPNDEGRYTFGASCVTACPYNYLSTDVGSCTIYC 315  
 275 LHPNHSIGICELCPALVTYNTDFFESMPNDEGRYTFGASCVTACPYNYLSTDVGSASIT- 333  
 316 PLHNOEVTAEADGTORCEKSKPCARVCYGLGMEHLREVRVAVTSANIOEPAGCKKIFGSLA 375  
 334 ----- 333  
 376 FLPESTGDPASNTAPLQEPOLQVFTLEITGYLYISAMPDSLPLSVFQNLQVIRGR 435  
 334 -----NEFKL----- 338  
 436 LHNAGVSLTLOGIGISWLGSLRLSRLSGSLALIHNTHLICFVHTVPWDLFRPHQALLH 495  
 339 ----- 338  
 496 TANRPEDECVGEGIACHQIACARHCWGPPTOCVNCISOFLRGQECVEECRVLOGLPREYV 555  
 339 ----- 338  
 556 NARHCLPCHPECOPONGSVTCFGRPADQCVACAHYKDPFCVYARCPGSKVYKPDLSYMPIK 615  
 339 ----- 338  
 616 FPDEBACQPCPTINCTHSCVDLDDKGCRAEORASPLTSONEDLGPASPLDSTFYRSLLED 675  
 339 ----- 338  
 676 DMWGLVDAEYLVPQGGFCPPDAPGAGMHHRRSSSTRSGGDLTLGLEPSEEAR 735  
 339 -----GAGGMHHRRSSSTRSGGDLTLGLEPSEEAR 372  
 736 RSPPLASSEGAGSDVFDGLGMAKGLQSLPTHDPSPLORYSEDPVPLPSETDGVAPL 795  
 373 RSPPLASSEGAGSDVFDGLGMAKGLQSLPTHDPSPLORYSEDPVPLPSETDGVAPL 432  
 796 TCSPOPEYVNOPDVPRQPPSPREGPLPARPAGATLERKATLSPGNGVYKDVFAFGAV 855  
 433 TCSPOPEYVNOPDVPRQPPSPREGPLPARPAGATLERKATLSPGNGVYKDVFAFGAV 492  
 856 ENPEYTLTPOGGAAPORHPRAPSPAFNDLYYWDOPPERGAPSPSTKGTPTAENPEYTL 915  
 493 ENPEYTLTPOGGAAPORHPRAPSPAFNDLYYWDOPPERGAPSPSTKGTPTAENPEYTL 552  
 916 DVP 918  
 553 DVP 555

Result 15  
 US-09-821-883-4  
 Sequence 4, Application US/09821883  
 Patent No. US20020061310A1  
 GENERAL INFORMATION:  
 APPLICANT: Vaid, Reiner  
 APPLICANT: Laidovic, Damir  
 APPLICANT: Graddis, Thomas

TITLE OF INVENTION: Compositions and Methods for Dendritic  
 TITLE OF INVENTION: Cell-Based Immunotherapy  
 FILE REFERENCE: 7636-0022.30  
 CURRENT APPLICATION NUMBER: US/09/821,883  
 CURRENT FILING DATE: 2001-03-30  
 PRIOR APPLICATION NUMBER: US 60/193,504  
 PRIOR FILING DATE: 2000-03-30  
 NUMBER OF SEQ ID NOS: 30  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 4  
 LENGTH: 697

TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: HER500\*-rsm-CSF construct  
 US-09-821-883-4

Query Match 50.9%; Score 2585; DB 10; Length 697;  
 Best Local Similarity 56.7%; Pred. No. 7,2e-142;  
 Matches 512; Conservative 3; Mismatches 6; Indels 382; Gaps 2;

16 LPPGASTVCTGTDKMLRLPASPEHMLRLHLYGGCOVVOGNLELTLPNTASISFQ 75  
 35 LARGAASVCTGTDKMLRLPASPEHMLRLHLYGGCOVVOGNLELTLPNTASISFQ 94  
 76 DIOEVGYVLIANOVROYPLQRLRIYRGTOLEFEDNYALAVLDNGPPLNNTPVYTGASPG 135  
 95 DIOEVGYVLIANOVROYPLQRLRIYRGTOLEFEDNYALAVLDNGPPLNNTPVYTGASPG 154  
 136 GLREIOLRLSTELIKGVLIOIRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPC 195  
 155 GLREIOLRLSTELIKGVLIOIRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPC 214  
 196 SPWCKSGRWGSSSEDCOSLTRVCAGGACRCKGRLPTDCHEQCAAGCTGPKHSCLAC 255  
 215 SPWCKSGRWGSSSEDCOSLTRVCAGGACRCKGRLPTDCHEQCAAGCTGPKHSCLAC 274  
 256 LHPNHSIGICELCPALVTYNTDFFESMPNDEGRYTFGASCVTACPYNYLSTDVGSCTIYC 315  
 275 LHPNHSIGICELCPALVTYNTDFFESMPNDEGRYTFGASCVTACPYNYLSTDVGSASIT- 333  
 316 PLHNOEVTAEADGTORCEKSKPCARVCYGLGMEHLREVRVAVTSANIOEPAGCKKIFGSLA 375  
 334 ----- 333  
 376 FLPESTGDPASNTAPLQEPOLQVFTLEITGYLYISAMPDSLPLSVFQNLQVIRGR 435  
 334 -----NEFKL----- 338  
 436 LHNAGVSLTLOGIGISWLGSLRLSRLSGSLALIHNTHLICFVHTVPWDLFRPHQALLH 495  
 339 ----- 338  
 496 TANRPEDECVGEGIACHQIACARHCWGPPTOCVNCISOFLRGQECVEECRVLOGLPREYV 555  
 339 ----- 338  
 556 NARHCLPCHPECOPONGSVTCFGRPADQCVACAHYKDPFCVYARCPGSKVYKPDLSYMPIK 615  
 339 ----- 338  
 616 FPDEBACQPCPTINCTHSCVDLDDKGCRAEORASPLTSONEDLGPASPLDSTFYRSLLED 675  
 339 ----- 338

Qy	676	DDMDGLVDAEEYLVPQGGFFCPDPAAPGAGVHHRRHSSSTRBGGGLTLGLEPSSSEAP	735
Dd	339	-----GAGGVHHRRHSSSTRBGGGLTLGLEPSSSEAP	372
Qy	736	RSPLAPSEGAGSDVFDGDLGGAAGKGLQSLPTHDPSPLOQYSEDPTVPLPSETDGYAPL	795
Dd	373	RSPLAPSEGAGSDVFDGDLGGAAGKGLQSLPTHDPSPLOQYSEDPTVPLPSETDGYAPL	432
Qy	796	TCSPOPEYVNDPVRPOPSPRREGPLLAARPAGATLERTLSLPGKGVYKDVFAFGAV	855
Dd	433	TCSPOPEYVNDPVRPOPSPRREGPLLAARPAGATLERTLSLPGKGVYKDVFAFGAV	492
Qy	856	ENPEYLTPQGGADQPHPPPAFSAFDNLVYMDDPERCAPSTFKGTPTAENPEYLG	915
Dd	493	ENPEYLTPQGGADQPHPPPAFSAFDNLVYMDDPERCAPSTFKGTPTAENPEYLG	552
Qy	916	DVP 918	
Dd	553	DVP 555	

Job completed: January 13, 2003, 14:50:16  
Time : 22.7057 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2003, 14:44:55 : Search time 22.5383 Seconds  
(without alignments)  
3919.881 Million cell updates/sec

Title: US-09-854-356-6  
Perfect score: 5078  
Sequence: 1 MELALICRMWGLLALLPPGA.....TFKGNPTAENPEYGLDVPV 919

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%  
Listing first 405 summaries

Database : PIR\_73:\*

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4900	96.5	1255	1 A24571	protein-tyrosine k
2	4134	81.4	1260	1 TVKTNV	protein-tyrosine k
3	4122.5	81.2	1254	2 I48161	p-185 precursor
4	1645	32.4	1210	1 GQHUE	epidermal growth f
5	1612	31.7	1210	2 A53183	epidermal growth f
6	1592.5	31.4	1223	1 TVCHLV	epidermal growth f
7	1536	30.2	644	2 A36325	epidermal growth f
8	1528.5	30.1	1308	2 A47253	epidermal growth f
9	1435	28.3	1342	2 A36223	kinase-related tra
10	1351.5	26.6	1339	2 JC4387	epidermal growth f
11	1313.5	25.9	1166	1 S06142	protein-tyrosine k
12	1155	22.7	527	2 A42032	epidermal growth f
13	1014.5	20.0	843	2 A27131	epidermal growth f
14	718.5	14.1	1323	2 E88257	protein let-23 [tm
15	718.5	14.1	1374	2 S70712	protein-tyrosine k
16	693	13.6	1330	1 GQFE	epidermal growth f
17	669.5	13.2	1369	2 S70713	protein-tyrosine k
18	621	12.2	1717	1 A45558	epidermal growth f
19	384	7.6	1363	2 T43220	insulin-like growt
20	376.5	7.4	1300	2 A36502	insulin receptor-t
21	370	7.3	1382	1 INHUR	insulin receptor p
22	363.5	7.2	366	2 D45558	epidermal growth f
23	362.5	7.1	1383	2 A36080	insulin receptor p
24	361	7.1	1372	2 A34157	insulin receptor p
25	353	7.0	1477	2 T18534	protein-tyrosine k
26	348.5	6.9	540	2 B47417	insulin receptor-t
27	338	6.7	1268	2 B36502	insulin receptor-t
28	331	6.5	333	2 B45558	epidermal growth f
29	331	6.5	342	2 C45558	epidermal growth f

30	326	6.4	1390	2 T30346	insulin receptor -
31	324	6.4	1607	2 T43212	insulin-like growt
32	316	6.2	1371	2 A33837	insulin-like growt
33	311.5	6.1	1367	1 IGHUR1	insulin-like growt
34	307	6.0	2101	2 S57245	insulin receptor (
35	307	6.0	2148	1 A56081	insulin receptor k
36	299.5	5.9	698	1 TVFVLV	protein-tyrosine k
37	267.5	5.3	1846	2 T42047	insulin receptor h
38	254.5	5.0	604	1 TVYUHH	protein-tyrosine k
39	244.5	4.8	1548	2 S34583	serine proteinase
40	240	4.7	1299	2 T43251	furin (EC 3.4.21.7
41	235	4.6	329	2 A48805	insulin-like growt
42	226.5	4.5	183	2 JH0803	tyrosine kinase re
43	208	4.1	1680	2 A43434	furin (EC 3.4.21.7
44	200	3.9	545	2 S00727	kinase-related tra
45	198.5	3.9	962	2 JC5571	subtilisin-like pr

## ALIGNMENTS

RESULT 1  
A24571  
protein-tyrosine kinase (EC 2.7.1.112) erbB2 precursor - human  
N:Alternate names: c-erb-B-2 protein precursor; kinase-related transforming protein e  
C:Species: Homo sapiens (man)  
C:Date: 25-Oct-1987 #sequence revision 06-Dec-1996 #text change 11-Jun-1999  
C:Accession: A24571; A25491; A44188; B44188; I59509; I57622  
R:Yamamoto, T.; Ikawa, S.; Akiyama, T.; Semba, K.; Nomura, N.; Miyajima, N.; Saito, T  
Nature 319, 230-234, 1986  
A>Title: Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth  
A:Reference number: A24571; MUID:86118663; PMID:3003577  
A:Accession: A24571  
A:Molecule type: mRNA  
A:Residues: 1-1255 <YAM>  
A:Cross-References: GB:X03363; NID:931197; PIDN:CAA27060.1; PID:931198  
R:Semba, K.; Kamata, N.; Toyoshima, K.; Yamamoto, T.  
Proc Natl. Acad. Sci. U.S.A. 82, 6497-6501, 1985  
A>Title: A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-erbB-1/epid  
A:Reference number: A25491; MUID:86016729; PMID:2995967  
A:Accession: A25491  
A:Molecule type: DNA  
A:Residues: 737-1031 <SEM>  
A:Cross-References: GB:M11767; NID:9182163; PIDN:AAA35808.1; PID:9553282  
R:Cousens, L.; Yang-Peng, T.L.; Liao, Y.C.; Chen, E.; Gray, A.; McGrath, J.; Seeburg  
Science 230, 1132-1139, 1985  
A>Title: Tyrosine kinase receptor with extensive homology to EGF receptor shares chro  
A:Reference number: A44188; MUID:86070181; PMID:2999974  
A:Accession: A44188  
A:Molecule type: DNA  
A:Residues: 740-910 <COU>  
A:Cross-References: GB:M12036; NID:9183988; PIDN:AAA35978.1; PID:9183989  
A:Accession: B44188  
A:Molecule type: mRNA  
A:Residues: 1-517; 'RALT', 522, 'S', 524-654, 'V', 656-1169, 'A', 1171-1255 <COU2>  
A:Cross-References: GB:M11730; NID:9183986  
R:King, C.R.; Kraus, M.H.; Aaronson, S.A.  
Science 229, 974-976, 1985  
A>Title: Amplification of a novel v-erbB-related gene in a human mammary carcinoma.  
A:Reference number: I59509; MUID:85272597; PMID:2992089  
A:Accession: I59509  
A:Molecule type: DNA  
A>Status: translated from GB/EMBL/DBJ  
A:Residues: 832-909 <REX>  
A:Cross-References: GB:L29395; NID:9459807; PIDN:AAA35809.1; PID:9459808  
R:Tal, M.; King, C.R.; Kraus, M.H.; Ullrich, A.; Schlessinger, J.; Givol, D.  
Mol. Cell. Biol. 7, 2597-2601, 1987  
A>Title: Human HER2 (neu) promoter: evidence for multiple mechanisms for transcriptio  
A:Reference number: I57622; MUID:87286898; PMID:3039351  
A:Accession: I57622  
A:Molecule type: DNA  
A>Status: translated from GB/EMBL/DBJ  
A:Residues: 1-191 <TAU>



A:Cross-references: GB:M16792; NID:g183983; PIDN:AAA58637.1; PID:g9553332  
 C:Comment: Amplification and overexpression of this erbB-related gene occurs in about 30  
 C:Genetics:  
 A:Gene: GDB:ERB2; NGL; NEU; HER-2  
 A:Cross-references: GDB:120613; OMIM:164870  
 A:Map position: 17q21.1-17q21.1  
 A:Introns: 25/1; 75/3; 147/1; 883/3  
 A>Note: the list of introns is incomplete  
 C:Function:  
 A:Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP  
 C:Superfamily: epidermal growth factor receptor; protein kinase homology  
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho  
 kinase  
 F:1-21/Domain: signal sequence #status predicted <SIG>  
 F:22-155/Product: protein-tyrosine kinase erbB2 #status predicted <MAT>  
 F:22-653/Domain: extracellular #status predicted <EXT>  
 F:70-304/Domain: EGF receptor extracellular domain repeat <EB1>  
 F:395-605/Domain: EGF receptor extracellular domain repeat <EB2>  
 F:654-675/Domain: transmembrane #status predicted <TM>  
 F:676-1253/Domain: intracellular #status predicted <INT>  
 F:718-983/Domain: protein kinase homology <KIN>  
 F:726-734/Region: protein kinase ATP-binding motif  
 F:68,124,187,259,530,571,629/Binding site: carbohydrate (Asn) (covalent) #status predict  
 F:686/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted  
 F:759/Active site: lys #status predicted  
 F:1139,1221,1222,1248/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation)

Query Match 96.5%; Score 4900; DB 1; Length 1255;  
 Best local Similarity 73.2%; Pred. No. 3.3e-263;  
 Matches 919; Conservative 0; Mismatches 0; Indels 336; Gaps 1;

OY . 1 METALCRWGLLALLPRGAASVCTGTDMKRLRSPETHDMLRLHYOGCVVGNL 60  
 Db 1 MELALCRWGLLALLPRGAASVCTGTDMKRLRSPETHDMLRLHYOGCVVGNL 60  
 OY 61 ELYTLPTNASISFTQDIOEVGYVLIANQVROYPLQRLTVRGSTOLFEDNYALAVDNG 120  
 Db 61 ELYTLPTNASISFTQDIOEVGYVLIANQVROYPLQRLTVRGSTOLFEDNYALAVDNG 120  
 OY 121 DPLNNTPTVYGASPGGLRELOLRSLTEILKGVLIQRNPOLCYDTILMKDIFKHNQLA 180  
 Db 121 DPLNNTPTVYGASPGGLRELOLRSLTEILKGVLIQRNPOLCYDTILMKDIFKHNQLA 180  
 OY 181 LTLIDTNSRACHPCSPCKSGRCWGESSEDCQSLFTVTCAGGACRCKGRLPTDCHEQC 240  
 Db 181 LTLIDTNSRACHPCSPCKSGRCWGESSEDCQSLFTVTCAGGACRCKGRLPTDCHEQC 240  
 OY 241 AAGCTGPRHSDCLAFHNHSGICELHCPALVTYNTDTFESMPNEGXYTGASCVTACP 300  
 Db 241 AAGCTGPRHSDCLAFHNHSGICELHCPALVTYNTDTFESMPNEGXYTGASCVTACP 300  
 OY 301 YNYLSTVGSCSTIVCPILHNQVTAEDGTQRCCKSPCARCYGIGMHLREVRAYTSAN 360  
 Db 301 YNYLSTVGSCSTIVCPILHNQVTAEDGTQRCCKSPCARCYGIGMHLREVRAYTSAN 360  
 OY 361 IOEFAGCKKIFGSLAFPEFSGDPASNTAPLQEPQLOVFTEETGYLISAMPDSL 420  
 Db 361 IOEFAGCKKIFGSLAFPEFSGDPASNTAPLQEPQLOVFTEETGYLISAMPDSL 420  
 OY 421 DLSFQNLQVYRGRILNNGASITLQIGISWLGIRSLRELGSGALLNHNTHLCFVTV 480  
 Db 421 DLSFQNLQVYRGRILNNGASITLQIGISWLGIRSLRELGSGALLNHNTHLCFVTV 480  
 OY 481 PMDOLFNRPHOALLHTANRPEDECVGEGLACHOLCARGCHGPGPTQCVNCSQFLRGQEC 540  
 Db 481 PMDOLFNRPHOALLHTANRPEDECVGEGLACHOLCARGCHGPGPTQCVNCSQFLRGQEC 540  
 OY 541 VEECRVLQGLRREYVNAHRCPCHECOPONGSVTCFSPREADQCVACAHYKDRPFCVARC 600  
 Db 541 VEECRVLQGLRREYVNAHRCPCHECOPONGSVTCFSPREADQCVACAHYKDRPFCVARC 600  
 OY 601 PSGVKPDLSTYPIWKFPDEBACOPCPINCTHSCVDLDDKCPAQRASPLTS----- 653  
 Db 601 PSGVKPDLSTYPIWKFPDEBACOPCPINCTHSCVDLDDKCPAQRASPLTS----- 653

Db 601 PSGVKPDLSTYPIWKFPDEBACOPCPINCTHSCVDLDDKCPAQRASPLTSISAVYG 660  
 OY 654 ----- 653  
 Db 661 ILLVVLGVVFGILIKRRQOKIRKRYTMRRLQETELVEPLTPSGAMNQAMRIKETEL 720  
 OY 654 ----- 653  
 Db 721 RKKVKLGSAGFTVYKGIWIPDGENVKIPVAIKYLRNTPSKANKETLDEAYVAGVGP 780  
 OY 654 ----- 653  
 Db 781 YVSRLLGICLTSTVQVLTQMLPFGCLLDHYRENKRGSDILNMCQIAKMSYLEDVR 840  
 OY 654 ----- 653  
 Db 841 LVHRDLAARNVLYKSPNHVKITDGLARLDDIDETEXHADGKVPIMMALESILRRFT 900  
 OY 654 ----- 653  
 Db 901 HQSDWSYGVTVWELMTFGAKPYDGIIPAREIPLLEKGERLPQPICTIDVYIMVKCM 960  
 OY 654 ----- 653  
 Db 961 IDSCRRFRFLVSEFSRMARDPQRFYVIONEDLGRASPILDSTTYRSLLEDMDKGLVDA 1020  
 OY 685 EELYLPQGGFPCPDPAAGAGVHHRHRSSTRSGGDLTLGLPSEEARPSPLASEG 744  
 Db 1021 EELYLPQGGFPCPDPAAGAGVHHRHRSSTRSGGDLTLGLPSEEARPSPLASEG 1080  
 OY 745 AGSVFPGDGLMGAAKGLQSLPTHDPSPLQKYSDDPVPLPSENDGYVAPLTCSPQPEYV 804  
 Db 1081 AGSVFPGDGLMGAAKGLQSLPTHDPSPLQKYSDDPVPLPSENDGYVAPLTCSPQPEYV 1140  
 OY 805 NOPDVRRQPPSPREGPPLPAARPAATLERPTLSPKNGVYKDVFAFGAENPEYLTPO 864  
 Db 1141 NOPDVRRQPPSPREGPPLPAARPAATLERPTLSPKNGVYKDVFAFGAENPEYLTPO 1200  
 OY\* 865 GGAAPQHPAPPAFAFDNLTYWDODPPERGAPSTFKGTPTAENPEYLGIDPV 919  
 Db 1201 GGAAPQHPAPPAFAFDNLTYWDODPPERGAPSTFKGTPTAENPEYLGIDPV 1255

RESULT 2  
 TVRKNL  
 protein-tyrosine kinase (EC 2.7.1.112) neu precursor - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 11-Jun-1999  
 C:Accession: A24562; A61204  
 R:Barigman, C.I.; Hung, M.C.; Weinberg, R.A.  
 Nature 319, 226-230, 1986  
 A:Title: The neu oncogene encodes an epidermal growth factor receptor-related protein  
 A:Reference number: A24562; MUID:86118662; PMID:3945311  
 A:Accession: A24562  
 A:Molecule type: mRNA  
 A:Residues: 1-1260 <BAR>  
 A:Cross-references: EMBL:X03362; NID:g56745; PIDN:CAA27059.1; PID:g56746  
 R:Masui, T.; Mann, A.M.; Maattee, T.L.; Garland, E.M.; Okamura, T.; Smith, R.A.; Cole  
 Carcinogenesis 12, 1975-1978, 1991  
 A:Title: Direct DNA sequencing of the rat neu oncogene transmembrane domain reveals a  
 2-thiazolyl]formamide or N-methyl-N-nitrosourea  
 A:Reference number: A61204; MUID:92035293; PMID:1682063  
 A:Accession: A61204  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 637-663, 'V', 665-702 <MAS>  
 A>Note: authors translated the codon GCA for residue 25 as Val  
 C:Genetics:  
 A:Gene: neu  
 C:Superfamily: epidermal growth factor receptor; protein kinase homology  
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phos  
 F:1-19/Domain: signal sequence #status predicted <SIG>  
 F:20-1260/Product: protein-tyrosine kinase neu #status predicted <MAT>

F:658-680/Domain: transmembrane #status predicted <TMN>

F:723-988/Domain: protein kinase homology <KIN>

F:731-739/Region: protein kinase ATP-binding motif

F:711,131,163,535,576,634/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:691/Binding site: phosphate (Thr) (covalent) #status predicted

F:758/Active site: Lys #status predicted

F:882,1227,1253/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 81.4%; Score 4134; DB 1; Length 1260;

Best Local Similarity 62.1%; Pred. No. 7.1e-221;

Matches 781; Conservative 42; Mismatches 96; Indels 338; Gaps 3;

```

OY 1 METLAALCRMGILLALLPFGAASVQVCTGTDMLRLPASPTHTDMLRHLYOGGQVVOGNTL 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4 MELAAACRMGFLALLPFGIAGTQVCTGTDMLRLPASPTHTDMLRHLYOGGQVVOGNTL 63
OY 61 ELYTLPNASTLFLQDIOEVGYVLIANNOVROVPLQRLRIVRTQLEFEDNYALAVLDNG 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 64 ELYTVPANASTLFLQDIOEVGYVLIANNOVKRVPLORLRIVRTQLEFEDKYALAVLDNR 123
OY 121 DPLNNTPTVY-GASPGGLRELQRLSLTEILKGVLIQNRQOLCYQDTILMKDIFHKNNQL 179
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 124 DPLDNVAASVTPGRTPBELRLQRLSLTEILKGVLIQNRQOLCYQDMVYLRKNNQL 183
OY 180 ALTLIDNRSRACHPCSPMCKSGSRMGESSEDCOSLRTVACAGCARGKPLPTDCCHQ 239
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 184 APVDIDNRSRACHPCSPMCKSGSRMGESSEDCOSLRTVACAGCARGKPLPTDCCHQ 243
OY 240 CAAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDPEFSPNBPGRYTFGASCVTAC 299
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 244 CAAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDPEFSPNBPGRYTFGASCVTTC 303
OY 300 PNYVLTSDVSCVLYNGLHNOEVTABDGTQRCCKSPCARVCYGLGMEHLREVAVTSA 359
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 304 PNYVLTSEVSCVLYNGLHNOEVTABDGTQRCCKSPCARVCYGLGMEHLRGAALTS 363
OY 360 NIOFAGCKKIFGSLAFLPSPFGDPPASNTAPLOPEQLQVETLEITGTYLISAMPDSL 419
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 364 NIOFAGCKKIFGSLAFLPSPFGDPPASNTAPLOPEQLQVETLEITGTYLISAMPDSL 423
OY 420 PDLVSFQNLQVIRGLIHNGAYSLLTQGLGISWLGSLRELSSGALLHHNTLFCVHT 479
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 424 RDLVSFQNLQVIRGLIHNGAYSLLTQGLGISWLGSLRELSSGALLHHNTLFCVHT 483
OY 480 VPMDOLEFRNHQALHNGANPEDE-CYBEGSLACHQICARHCHMGPGTQCVNCSQFLRG 538
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 484 VPMDOLEFRNHQALHNGANPEDE-CYBEGSLACHQICARHCHMGPGTQCVNCSQFLRG 543
OY 539 ECVEBCHVLOGLPREVYNARHCLPCHEPCOPONGSVTCFGEADQCYAACHYKDPFCVA 598
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 544 ECVEBCHVLOGLPREVYNARHCLPCHEPCOPONGSVTCFGEADQCYAACHYKDPFCVA 603
OY 599 RCBGKVPDLSTYPIWKFPEBEGACOPCPINCTHSCVDLDDKGCAPABQASPLT----- 652
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 604 RCBGKVPDLSTYPIWKFPEBEGACOPCPINCTHSCVDLDDKGCAPABQASPLT----- 663
OY 653 ----- 652
Db 664 EGVLLFLIVVVVGLIKRRRQRIKRYTMRLLQETELVEPLTPSGAMPNQAQMLIKET 723
OY 653 ----- 652
Db 724 ELKRVKYLSSGAGTYVKGIMIPDGENVKIPVAIKVIRENTSPKANKEILDEAYVMAVG 783
OY 653 ----- 652
Db 784 SPVYSRLIGLITSTVQLVTLMPYGLDHDVREHNRGLSGODLMMCVQIAKMSYLE 843
OY 653 ----- 652
Db 844 VRLVHRDLAARNVLVSPNHVKTITDFGLARLDIDETEHADGKVPKIMMALESILRR 903
OY 653 ----- 652

```

```

Db 904 FTHOSDVWSYGVYWEIMLTGAKRYDGIAPRELIDLEKBERLPQPICTIDVYMIWVC 963
OY 653 -----SQNEDLCASPDLSTFYRSLLEDMDMDLV 682
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 964 WMIDSECRPRFRELSEFSSKWARDPQRFVVIYQNEIDLSPSPMDSTFYRSLLEDMDMDLV 1023
OY 683 DAERYLVPOQGFPCPPDAPAGAGVHHRRSSSTRSGGDLTGLSESEAPRSLAPS 742
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1024 DAERYLVPOQGFSPDTPPTGTAAHRRHSSSTRSGGDLTGLSESEAPRSLAPS 1083
OY 743 EGASDVFDGDLGKAQGLQSLPTHDPSLQORSEPTVPLPSETDGYAAPLTCSPQPE 802
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1084 EGASDVFDGDLGKAQGLQSLPTHDPSLQORSEPTVPLPSETDGYAAPLTCSPQPE 1143
OY 803 YVNOQVROPSPREBGLPLAARPAATLERPKTLSPGKNGVYKDYAFGAVENPEYLT 862
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1144 YVNOQVROPSPREBGLPLAARPAATLERPKTLSPGKNGVYKDYAFGAVENPEYLT 1203
OY 863 POGGAAPRRPFPSPAFNDLVYWDOPPREBAPSPSTFGTPTAENPEYLTGLDVPY 919
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1204 PRESTASPPHSPAFSPAFNDLVYWDQNSSEQGPSPNFEGTPTAENPEYLTGLDVPY 1260

```

# RESULT 3

I48161

P-185 precursor - golden hamster

C:Species: Mesocricetus auratus (golden hamster)

C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 18-Jun-1999

C:Accession: I48161

R:Nakamura, T.; Ushijima, T.; Ishizaka, Y.; Nagao, M.; Arai, M.; Yamazaki, Y.; Ishika

A>Title: Cloning and activation of the Syrian hamster neu proto-oncogene.

A:Reference number: I48161; MID:94193007; PMID:7908275

A:Accession: I48161

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1254 <RES>

A:Cross-references: GB:D16295; NID:9493236; PIDN:BA03801.1; PID:9747595

C:Genetics:

A:Gene: neu

C:Superfamily: epidermal growth factor receptor; protein kinase homology

C:Keywords: ATP

F:718-983/Domain: protein kinase homology <KIN>

F:726-734/Region: protein kinase ATP-binding motif

Query Match 81.2%; Score 4122.5; DB 2; Length 1254;

Best Local Similarity 61.8%; Pred. No. 3.1e-220;

Matches 776; Conservative 49; Mismatches 93; Indels 337; Gaps 2;

```

OY 1 METLAALCRMGILLALLPFGAASVQVCTGTDMLRLPASPTHTDMLRHLYOGGQVVOGNTL 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MELAAACRMGILLALLPFGASGTVCTGTDMLRLPASPTHTDMLRHLYOGGQVVOGNTL 60
OY 61 ELYTLPNASTLFLQDIOEVGYVLIANNOVROVPLQRLRIVRTQLEFEDNYALAVLDNG 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 ELYTLPNASTLFLQDIOEVGYVLIANNOVROVPLQRLRIVRTQLEFEDNYALAVLDNR 120
OY 121 DPLNNTPTVY-GASPGGLRELQRLSLTEILKGVLIQNRQOLCYQDTILMKDIFHKNNQL 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 DPLDNVAASVTPGRTPBELRLQRLSLTEILKGVLIQNRQOLCYQDMVYLRKNNQL 180
OY 181 LTLIDNRSRACHPCSPMCKSGSRMGESSEDCOSLRTVACAGCARGKPLPTDCCHQ 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 PVDIDNRSRACHPCSPMCKSGSRMGESSEDCOSLRTVACAGCARGKPLPTDCCHQ 240
OY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDPEFSPNBPGRYTFGASCVTAC 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDPEFSPNBPGRYTFGASCVTTC 300
OY 301 YNLTSTDVGSCTLVCPILHNOEVTABDGTQRCCKSPCARVCYGLGMEHLREVAVTSA 360
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 YNLTSTFVGSCTLVCPILHNOEVTABDGTQRCCKSPCARVCYGLGMEHLRGAALTSAN 360

```



A: Note: the EGF receptor (and other tyrosine kinases) can nick double-stranded DNA  
 R: Chen, W.S.; Lazar, C.S.; Lund, K.A.; Welsh, J.B.; Chang, C.P.; Walton, G.M.; Der, C.J.  
 Cell 59, 33-43, 1989  
 A: Title: Functional independence of the epidermal growth factor receptor from a domain  
 A: Reference number: A33331; MID: 90003233; PMID: 2790960  
 A: Contents: annotation; internalization signal  
 C: Comment: Binding of EGF to the receptor leads to internalization of the EGF-receptor  
 C: Genes:  
 A: Gene: GDB: EGRF  
 A: Cross-references: GDB: 120610; OMIM: 131550  
 A: Map position: 7p12.3-7p12.1  
 C: Superfamily: epidermal growth factor receptor; protein kinase homology  
 C: Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho  
 F: 1-24/Domain: signal sequence #status predicted <SIS>  
 F: 25-1210/Product: EGF receptor #status predicted <MAT>  
 F: 25-645/Domain: extracellular #status predicted <EXT>  
 F: 75-300/Domain: EGF receptor extracellular domain repeat <EE1>  
 F: 390-600/Domain: EGF receptor extracellular domain repeat <EE2>  
 F: 646-668/Domain: transmembrane #status predicted <TM>  
 F: 669-1210/Domain: intracellular #status predicted <INT>  
 F: 710-925/Domain: protein kinase homology <KIN>  
 F: 718-726/Region: protein kinase ATP-binding motif  
 F: 999-1046/Region: coated-pit mediated internalization signal  
 F: 1047-1210/Region: inhibitory  
 F: 128, 175, 352, 413, 444, 528, 603/Binding site: carbohydrate (Asn) (covalent) #status predic  
 F: 745/Active site: Lys #status experimental

Query Match 32.4%; Score 1645; DB 1; Length 1210;  
 Best Local Similarity 29.6%; Pred. No. 2,4e-83;  
 Matches 374; Conservative 140; Mismatches 309; Indels 442; Gaps 22;

11 LLLALLPQGA--STOVCTGDMKLRLPASPEHMDLRHLGGCOVVOGNTLETLPTN 68  
 14 LLLALCPASALPEKKYKCOGTSNKLTLQGFEDHFLSLQRMFNCEVGLNLETTYORN 73  
 69 ASLFLDIDQVQGYVLIANQVROYPLQRLIVRGTOLEEDNYALVLDNGPLNNTTP 128  
 74 YDLSPFLKIOEVAGYVLIANTVERIPLENLQIRGMYYENSVLAVLSNYD----- 126  
 129 VFGASPGELRELQRLSTTEILKGVLIQNPOLCYODTILMKDIFKNNOLATLIDTNR 188  
 127 ---ANKTGLKELPMRNLQELIHGAVRFSNNPALSISQWRDIVSDFLSNMSDRPN 183  
 189 SRACHPCSPKCSRCMGESSEDCQSITRTVCAGGCA-RKKGPLPTCCHEGCAAGTGP 247  
 184 LGSQCQKCDPSCPGSCMGAGEENCQKLTLLCAQCCSGRCGRKSPSCCHNCAAGCTGP 243  
 248 KHSDCLACTLHFNHSGICELCPALVTYNTDTFESMPREGRTYFGASCVTACPYNYLSTD 307  
 244 RESDCLVCRKFRDEATCKDPCPLMLXNPTTYOMDVNPEKYSFGATCVAKCPRNYVTD 303  
 308 VGSCTLVCPILHNOEYAEIDGTQRCCKSCPKARVCYGLGMEHLREVRVAVTSANIQEPAGC 367  
 304 HGSCVRCAGADSYEM-EEDDVRRKCKCEGRCRVCNGIGIEPFRDSLINATNKIKHKNC 362  
 368 KRIFGSLAFPEPSFDGDSNTPARLPQQLQVETLEITGYLXISMPSLDLSTFQN 427  
 363 TTSISGDHLPLVAFRGDSFHTPPLDQELDKTKVKEITGFLLIQAMPENRNDLHAFEN 422  
 428 LQVIRGILHNGAVSLTLQGLISGLRLSRLREISGLALHNNTHLCFVHTVPMQDLFR 487  
 423 LEIRGTRKHGQPSLAVVSLNTSLGRSLKEISDGDYIISGNKNLCYANTINMKKLFEG 482  
 488 NPHQALLHTANRDEDECVEGGLACHOLCARGHCWGPPTQCVNCSOFLRQCEVEECRYL 547  
 483 TSQOKFTIISNRGNSCKATGVCVHALCSPEGCGPEPRDVCSCRVNSRGRECVDKCKLL 542  
 548 QGLPREVYVNAHRLPCHPEQOPQNSVTCGPRADOCVACAHKDDPPFCARPSGKPD 607  
 543 EGGPREVENSECIQCHPECLPQAMNNTCTGPRDNCIOCAHYIDGPHCVKTCPAGVME 602  
 608 LSYMPIKPFDEEGACOPCINCTHSCVDLDDGCPA----- 644

Db 603 NNTL-VMKYADAGVCHLCHPNCYGGCTGPGLGECPTNGKIPSIATGVALLLLVVA 661  
 Oy 645 -----EQRASPLTSONE----- 656  
 Db 662 LGIGLFMRHRHIVKRTLRRLLORELVEPLTPSGEAPNOALLRLKETEFKKIKVLGSG 721  
 Oy 657 ----- 656  
 Db 722 AFGTVYKGLMIPSEGEKKIVAIKELREATSPANKELIDEATVYMASVDNPHVCRLLGIC 781  
 Oy 657 ----- 656  
 Db 782 LSTVQLIQLMPFGCLLDVYREHKNGISQYLLNMCVQIAGKNYLEDRLVHRLAAR 841  
 Oy 657 -----DLG----- 659  
 Db 842 NVLKTPOHVKITDPFLAKLIGAEKEKYNAEGGKVPKMMALSTLHRYTHOSDWSYG 901  
 Oy 660 -----PA----- 661  
 Db 902 VYWEIMTFGSKPYDGIPIASEISILEKGERLPQPICTIDVYIMVWKMMIDADSHPKF 961  
 Oy 662 -----SPLDSTFYRSLLEDGMDGLVDABEYLVPOQ 692  
 Db 962 RELIEFSKMARDPQRYLVIGDERNHLPPTDSNFYRALMDEEDMDVDYDADEYLLPQ 1021  
 Oy 693 GFPCPPDPAGAGMVAHHRHSSSTRSGGDLTLGLEPSEEARPSPLASSEGASDVPDQ 752  
 Db 1022 GFF-----SSPSTSRPTLLSLSTASN--NS 1045  
 Oy 753 DLGGAAGKGLSLPTHDPSPLORESDEPTVPLPSET--DGVYAPLNCSPQPEVYVNDPDR 810  
 Db 1046 TVACIDRNGLSQCPICKEDSLQRYSSDPICALTEDSIDDTFL-----PYPETINO---- 1095  
 Oy 811 POPSPREPLPAPARPAATL-----ERPKTISPCKNGVYKDYAFGAVENPEYL-TPQ 864  
 Db 1096 -----SVF-KRPAQSQVNPYVHNPQPLNPAFSRDPHYQD--PSTANGNREYLNMTVQ 1143  
 Oy 865 GGAAPDHPPPASPADNLIYWDQ-----DP-----PENGAPSTKGTPTAE 908  
 Db 1144 -----PTCVNSTFSDSPAHMAOKGSHQISLDNPYQODFPPEKAPKIGIKGS-TAE 1193  
 Oy 909 NPEYL 913  
 Db 1194 NAEYL 1198

RESULT 5  
 A53183  
 epidermal growth factor receptor precursor - mouse  
 C: Species: Mus musculus (house mouse)  
 C: Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 18-Jun-1999  
 C: Accession: A53183; A43818; S24942; A28941; S45325; I49643  
 R: Luetteke, N.C.; Phillips, H.K.; Qiu, T.H.; Copeland, N.G.; Earp, H.S.; Jenkins, N.A  
 Genes Dev. 8, 399-413, 1994  
 A: Title: The mouse waved-2 phenotype results from a point mutation in the EGF receptor  
 A: Reference number: A53183; MID: 94170986; PMID: 8125255  
 A: Accession: A53183  
 A: Molecule type: mRNA  
 A: Residues: 1-1210 <LDB>  
 A: Cross-references: GB: U03425  
 R: Avioli, A.; Lax, I.; Ullrich, A.; Schlessinger, J.; Glyvol, D.; Morse, B.  
 Oncogene 6, 673-676, 1991  
 A: Title: Comparison of EGF receptor sequences as a guide to study the ligand binding  
 A: Reference number: A43818; MID: 91232866; PMID: 2030916  
 A: Accession: A43818  
 A: Molecule type: mRNA  
 A: Residues: 1-714 <AVI>  
 A: Cross-references: GB: X59698  
 R: Bisinger, D.P.; Serreio, G.  
 submitted to the EMBL Data Library, June 1992  
 A: Reference number: S24942  
 A: Accession: S24942



A:Accession: A27720  
A:Molecule type: mRNA  
A:Residues: 1-1223 <LAX>  
A:Cross-references: GB:M20386  
R:Nilsson, T.W.; Maroney, F.A.; Goodwin, R.G.; Rottman, F.M.; Crittenden, L.B.; Raines, M.  
Cell 41, 719-726, 1985  
A:Title: c-erbB activation in ALV-induced erythroblastosis: novel RNA processing and pro-  
A:Reference number: A00643; MUID:85328222; PMID:2988784  
A:Accession: A00643  
A:Molecule type: mRNA  
A:Residues: 585-1223 <NIID>  
A:Cross-references: GB:M10066  
C:Genetics:  
A:Gene: erbB  
C:Superfamily: epidermal growth factor receptor; protein kinase homology  
C:Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; growth factor  
specific protein kinase  
F:1-30/Domain: signal sequence #status predicted <SIG>  
F:31-1223/Product: epidermal growth factor receptor #status predicted <MAT>  
F:31-654/Domain: extracellular #status predicted <EXXT>  
F:81-307/Domain: EGF receptor extracellular domain repeat <EET>  
F:397-610/Domain: EGF receptor extracellular domain repeat <EET>  
F:653-677/Domain: transmembrane #status predicted <TM>  
F:678-1223/Domain: intracellular #status predicted <INT>  
F:719-984/Domain: protein kinase homology <KIN>  
F:727-735/Region: protein kinase ATP-binding motif  
F:136,202,280,361,370,422,575,580,615,635/Binding site: carbohydrate (Thr) (covalent) #  
F:192,650/Binding site: carbohydrate (Ser) (covalent) #status predicted  
F:687/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted  
F:754/Active site: Lys #status predicted  
F:1100,1183,1208/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #stat

Query Match 31.4%; Score 1592.5; DB 1; Length 1223;

Best Local Similarity 29.0%; Pred. No. 2e-80; Matches 375; Conservative 135; Mismatches 310; Indels 473; Gaps 25;

8 RMGLLALALPPGAA-----STOVCTGTMKRLRLPASPEHMLDMLHLYOGGVOVGNLE 61  
13 RGAAYLVLLLVALLGALSAVEKKVCGCTNNKLTOLGHEVDEHFTSLQRMNCEVLSNLE 72  
62 LTYLPTNASLSPLODIOEVGYVLLAHNOVROVPLQRLRIYNGTOLFEDNYALAVLNDG 121  
73 IYVEHNRLDTLKTIOEVAGYVLLAHNVAVLPLENLIQINGVLYDMSPLAVLSNH 132  
122 PLNNTPTVYGASBGRLREQLSLTEITLKGVLQORPOLCYQDTLTKMDIFHKNNQAL 181  
133 -NKKTO-----GLRELPMKRLSELINGVAKISNNPKLCMMDTVLAMDIDITSRK-PL 182  
182 TLID-TNRSRACHPCSPKCGSKRCMGESSEDCOSLRTVCAGGCA-RCKGRLPTDCCHQ 239  
183 TVLDFAASNLSSCPKCHPNCETEDHCWAGBQNCOTLTKVICAQCCSKRGKAVPSDCCHQ 242  
240 CAAGCGPKHSDCLALHFNHSGICELHCPALVYNTDTESMPNPEGRTFGASCVTAC 299  
243 CAAGCGPRESDCLACRKRFDATCKDCRPLVLYNPITYQMDVNEGKYSFGATCVRRC 302  
300 PNYLSTDVSGCTVLCPLHNOEYTAEDGTQREKSKPCARVCYGLGHEMLREAVATSA 359  
303 PNYVYVTDHSGVSCVSCNTDYEV-EENGVRKCKKCDGLSCVNGGIGELGIIISINAT 361  
360 NIOEFAGCKKIGSLAFLESPEDGPASNTAPLOPQLOVEFLEETITGYLISAMPDL 419  
362 NIDSKNCKKINGDYSILPVAFGLDAFTTLPDRKKLDVFTVAKISFGLIQAHPDNA 421  
420 PDLVFNQNLQVIRILHNGAVSLTQGLIGISWLGSLRLDGLSGLALIHNTHLCEVHT 479  
422 TDLVAFENLEIRGRTKQGYSLAVVNLKIQSLGRSLKEISDGDIALMKKNLCYADT 481  
480 VVMDQLFRPHALHTARPREDECVBELACHOLCARHCNGRPGTQCVNCSQPLRGOE 539  
482 MMRSLEFATOSOKRTIIQIRNRNCDTADRHVCDPLCSVDGCGPGEFHCFSRFSROKE 541  
540 CVEECRVLQGLPREVYNARHCLPCHPECOPNG---SVTCFGEADQCAVCAHYKDPKPC 596

542 CVQCNILGGEPRFERNDSKCLPCHSECLVQNSTAVNTQSGRGPDHCKAKAHFDGPHC 601  
597 VARCEGVAPDLSTYMIKFPDEGACQPCPINCHTSCVDLDDKCP----- 643  
602 VVACPGAVIGENDTL-VMKYADANAVCOLCHPNCFGCGPGEGCPNGSKTPSIAGVY 660  
644 -----AEGRASPLRSQNE----- 656  
661 GGLLCLVYVGLIGLYRRRHVYKRTLRLLQERLEVLTPSGEAPNQAHLRLIKETE 720  
657 ----- 656  
721 FKKVYLVSSGAGFTYKGLMIPBGEKVTIPVAKELREATSPRANKELIDPAVYMASVDN 780  
657 -----DLG----- 659  
841 RLVRHDLAARNVLKTPQHVKITDGLAKLGADEKEYHAEGKVPKWMALLESILHRTY 900  
660 -----PA----- 661  
901 THQSDVWSYGVYVWELMTFGSKPYDSIPASELSVLEKGERLPQPPICITDYVMIMVKW 960  
662 -----SPDSTFFYSRLDDEDMGDLV 682  
961 MIDADSRPFRELIAEFSKWARDPRVLYIQDERMHLRSPDSKFRYRLMEEDMEDIV 1020  
683 DAEYLYVPOQGFPCPRPAGAGGWNHHRSSSTNSGGDLTLGLEPSEEARPSPL--- 739  
1021 DDEYLYVPHQGF-----NSBST-----SRPTLSS 1046  
740 -APSEGASDVFDGLGKAAGLQSLPTHPSPLORYSEDPVPLPSET--DGVPAL 795  
1047 LSATSNKSNATCID-----RNGGHPYREDSEYQKISSPTGNFLESIDDGFL--- 1095  
796 TCSPOPEYVNOPDVPRQPPSPREGPLPAARPAATLERPKTLPKNGVYKDFV----- 849  
1096 -----PAPEYVNO--LMPKRS-----TAMVQNOYNNISLTAISK 1129  
850 -----AFEGAYENPEYLLPQGGAPQPPPAFPAFNLYWDO-----DPP 892  
1130 LPMDSRYQNSHSTAVDNPETL-----NTQOSPPLAKTFFESSPYWIOSGNHQINLNP 1181  
893 E-----RGAPSTFKGPTAENPEYLGIDVP 918  
1182 DYQDFLPMETKPNGLIKVPAENPEYLVANAP 1214

## RESULT 7

A36325 epidermal growth factor receptor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 25-Jan-1991 #sequence\_revision 25-Jan-1991 #text\_change 10-Oct-1997

C:Accession: A36325

R:Petich, L.A.; Harris, J.; Raymond, V.W.; Blasband, A.; Lee, D.C.; Barp, H.S.

Mol. Cell. Biol. 10, 2973-2982, 1990

A:Title: A truncated, secreted form of the epidermal growth factor receptor is encode

A:Reference number: A36325; MUID:90258888; PMID:2342466

A:Accession: A36325

A>Status: Preliminary

A:Molecule type: mRNA

A:Residues: 1-644 <PEY>

A:Cross-references: GB:M37394

C:Superfamily: epidermal growth factor receptor; protein kinase homology

C:Keywords: alternative splicing; ATP; growth factor receptor

Query Match 30.2%; Score 1536; DB 2; Length 644;  
Best Local Similarity 44.6%; Pred. No. 1.3e-77;  
Matches 287; Conservative 98; Mismatches 234; Indels 24; Gaps 8;



[illegible]

Db	8	WVVVSLVAACTVQPSDSQVCACTENKLSLSLDEQVRAIRKRYENCEVMNGLETT	67
Qy	65	LPTNASLFDLODIOEVGYVLIANOVROYPLQRLTIVBGTQLEFEDNALAVLNDGDLN	124
Db	68	IEHRDLSFLSVREYGVYVALVNOFRYLPLENLTIRIGTGLIEDRALAIFLYRRDG	127
Qy	125	NTTQVTCASPGCLAEJQBSLETILKGVLIQCNPOLCYQDPIILKDIFFHKNNQALALI	184
Db	128	NF-----GJQELGKLKLTETILNGGVYVQDNKFLCYADPIHODIVRNWPSNLTLY	178
Qy	185	DTNRSRACHPCSPKCKSRGMSSESDQSLRTYVAGGC-ARCKGPLPTDCHEQCAAG	243
Db	179	STNGSSGGRHKCTG-RQMGPEHNHCQTLRTTYVAEBCDRCYGRPYSDCHHECAGG	237
Qy	244	CTGRKSHDCLACLFHNSHGICELCPALVYTNTEFESMPNDEGRYTTGASCVTACPYN	303
Db	238	CSGKPDIDCFACMNFNDSGACTYQCQTFEYVNPPTTQLEHNHNAKTYTGAFVCWKCPNF	297
Qy	304	LSITVGSCTIIVCPAHNEVYAEADGTQRCCKSPCARVQYXGIMHLEAVRAYSANTQE	363
Db	298	V-VDSSTCVRACPSKMEV--EENGIMCKPCTDICKACDGGTGSLSMAQTVSSNDK	355
Qy	364	FAGCKITFGSLAFPESEFDDPASNAPLOPELOQYFTELEETITXYLIASMPDLPDS	423
Db	356	FINCTKINGNLITFYTHIOHDPYNALEADPEKLANFRKYRELTGFINQSPRMWDFS	415
Qy	424	VFQNLVYRGRILHNGAYSLTQLOGISWGLRSIRELSGLALIHNTLCFVHTBWD	483
Db	416	VFSNLVTIGGHVLTSGSLTLIKQGGITSLQFQSTKEISAGNIYITDINSNLCTYHTIMT	475
Qy	484	QLFNPHQALLHTANREDSCVEBGLAQHOLCARHGCMGPGYQVNCOSQPLRQGECEE	543
Db	476	TLESTINORLYIRDNRAENCTAEGMVCNHLSSDSCWGPGRDQCLSCRRSRKGTIES	535
Qy	544	CRVLQGLPREVYNAHRLCPHRECOR-ONGSVYCFEPEDQCVACAHKDPFYCARPS	602
Db	536	CNLDVGEREREENGSLCEVEDPQCEKMEBOLTGCHGRPDNCTKSHPRDGENCYEKPD	595
Qy	603	GVRQDLSMPIMKPFDEEGACQPCPCINCTHSCVDDLDKC-----PAEQR----	647
Db	596	GLQGANSE--TFKYADDPDRCHCHPCNCTQCGNTPSHDCIYVPTWGSHTLPOHARTPLI	653
Qy	648	-----ASPLTSQ-----	654
Db	654	AAGVIGLELIVIGLTFPAVYRRKRSIKKKRALRFLTELVEPLTPSGTAPNQAQRLIL	713
Qy	655	-----NEDLCP-----	660
Db	714	KETELKRRKVLGSGAFCTYKKGIMVEBGEVYKIPVAIKILNETTGPKANVERMBDALIMA	773
Qy	661	-----	660
Db	774	SMDPHLVRLIGVCLPTQLQVLMPHGCLLEYVHEKDNIGSOLLNMCQVAKGMW	833
Qy	661	-----	660
Db	834	LEBRRLVRLAARNVLYKSPNHVKITDFGLARILEGDEKEYNADGKMPIKMALBCTIH	893
Qy	661	-----	660
Db	894	YRKFTHOSDVMSYCVITWELMELTFGKRPYDGIPTREIPIDLLEKGERLPPOPICTIDVYWM	953
Qy	661	-----ASPLDSTFYSLLEDDM	678
Db	954	VCKMMIDADSPRKTELAEFSRMANDPORLYLVIOGDDMKLSPRSDSKFTQNLNLEDDL	1013
Qy	679	GDLVDAEYLVPOQGFCCPDPABGAGGVHHRHSSTRSGGGLTGLSPSEEBAPRS	737
Db	1014	EDMDAEYILVP-QAFNIPPP-----IYTSRAIDSNRS-----EIGHSPPATYPMSG	1061
Qy	738	-----PLAP--SEGAGSYVPRQDGLKMAKGLQSLPHND	769
Db	1062	NOFYVROGGFAAEQGVSPYRAPRTSTIPEPAVQAGTAIFIRDSOCNGTLKRPVAPHYOE	1121





Dd	171	GAEIYVKNNGANGACRPHCEVCKG - RCMGRPRDDCQILFTFTICAPQCGNCFGRPNRQCCD	229
OY	239	OCAAGCTGPKHSDCIACIAPHNHSGICELHCPALVYNTDTFESMPNPRGRYTFGASCYA	298
Dd	230	ECAGGCSPPODDCEACRRFNDSCAGVPRCEPIYKNTLFTQLEBNPHTKYQYGVCVAS	289
OY	239	CPYNTLSMDVSGCTIYVCLYHNEQVYAEADGTORCEKSCSPCARVCTYGLGMEHLREVRATVS	358
Dd	290	CPHNHV - VDQFFCVACRPRDKMEVD - KKGKMCCEPGGICAPACGTSG - -SRVQYUAS	345
OY	359	ANIOEFAACKIKFGSLAFLEPSFDGDPASNTAPLOPELOVETELETGYLYISAWPDS	418
Dd	346	SNIDGFVCTKILGMLDELITGLANDPWHKIPALDPEKLVNFRYTRREITGLYNIQSWPPI	405
OY	419	LPDLSVFONLQVIRRIIHNAGVS - LTLOGLSIWSIGRSIRLEISGALIIHHNTHLEFV	477
Dd	406	MHNESVFNULTTIGRSISLYNRGFSLLIMKNLVTSIGRSLSKEISAGVYIYSANQOLCYH	465
OY	478	HTVPMDQLEFRNHQALLHTA - NRPEDECVGEGSLACHOLCANRHCWMPGPTQCVNCSQFLR	536
Dd	466	HSIIMNTRILRGSEERLDIKTYDRPLGECLEABEGKVEDPLCSSGGCGMPAPGGCLSRANRS	525
OY	537	GOECVEECRVYLOGLEPREYVNAHNCILPCHPECOPONGSVTCFGRPEADQCVACAHYKDPFC	596
Dd	526	EGVCVTHCNFLTOGEEERFVHEAQCSPCHPECLPMEGTSTYNGSGSDACARCAHFRDQBHC	585
OY	597	VARCPSGVKPRDLSYPIKMFPEDEGACQPCPLNCTHSC	634
Dd	586	VNSCPHGILG - -AKGPYIKYPRDAQNECRPHENCTQGC	621

RESULT 11  
 S06142  
 protein-tyrosine kinase (EC 2.7.1.112) mrk-y precursor - southern platyfish  
 N:Alternate names: epidermal growth factor receptor homolog; kinase-related transforming  
 C:Species: Xiphophorus maculatus (southern platyfish)  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 18-Feb-2000  
 C:Accession: S06142; S13809  
 R:Witbrodt, J.; Adam, D.; Mallschek, B.; Maeuener, W.; Raulf, F.; Telling, A.; Roberts  
 Nature 341, 415-421, 1989  
 A>Title: Novel putative receptor tyrosine kinase encoded by the melanoma-inducing Tu loc  
 A:Reference number: S06142; MUID:90015140; PMID:2797166  
 A:Accession: S06142  
 A:Molecule type: DNA  
 A:Residues: 1-1166 <WIT>  
 A:Cross-references: EMBL:X16891; NID:965290; PIDN:CAA34770.1; PID:965291  
 R:Adam, D.; Maeuener, W.; Scharlt, M.  
 Oncogene 6, 73-80, 1991  
 A>Title: Transcriptional activation of the melanoma inducing Xmrk oncogene in Xiphophor  
 A:Reference number: S13807; MUID:91125882; PMID:11846957  
 A:Accession: S13809  
 A>Status: Preliminary; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 821-1025, 'N',1027-1098, 'A',1100-1166 <ADA>  
 A:Cross-references: EMBL:X56319; NID:965284; PIDN:CAA3763.1; PID:965285  
 C:Genetics:  
 A:Gene: mrk  
 A:Map position: y  
 A:Introns: 872/3; 898/1; 947/1; 979/3; 1025/3; 1056/1  
 C:Superfamily: epidermal growth factor receptor; protein kinase homology  
 C:Keywords: ATP; growth factor receptor; phosphotransferase; transmembrane protein; tyros  
 E:26-1166/Product: kinase-related transforming protein (Tu) #status predicted <MAT>  
 E:707-972/Domain: protein kinase homology <KIN>  
 E:715-723/Region: protein kinase ATP-binding motif

Query Match	25.9%	Score 1313.5	DB 1	Length 1166
Best Local Similarity	41.6%	Pred. No. 4.9e-65		
Matches 268	Conservative	93	Mismatches 259	Indels 25
				Gaps 12

```

Db      8 AALLD--LILLVLSIRCCSTDPDRKVCOGTSNQM---LDNHLYLKKAKKATSGCNYLLEN 62
OY      60 LELTYLPTNASLISFLQDIOEYGVYLLAHNOVROYVLORLRIVGTOLIFEDNYALAVLDN 119
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      63 LEITYOENODLSFIOSIOEYGVYLLAMNEVSTIPLVNLRLIRGONLYEGNFTLLWASN 122
OY      120 GDPUNNTTPYVGASGGIRELOLRSTETLLGCVLIQONPOLCYODPTLMDIFHKNNQL 179
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      123 YOK-PPSSP--DYVOVGKLOLQSLNLTETLLSGYKVASHNPLLCVETITMMDIVIDKTSNP 179
OY      180 ALTLITDRSRBACHPCSPMCKSGRCMCKSESSEDCOSLITFTVCAGGC-ARCKGPLTIDTCH 238
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      180 TMLNLIPIHAFEEQCKCHDGCVNCSGMARPGHCKKFTLLCAECCNNRGRGPKRIDCCNE 239
OY      239 OCAAGCTGPKHSDCLACLPHNSGICELHCPALVTYNTDTFESMPNBPGRYTFGASCVTA 298
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      240 HCAGGCTGPRATDCLACBDFNDGCTCKDTPPKPKIYDVSHQVVDNPIKTYTFEAGVCYE 299
OY      299 CPYNTLSTNDVSSCTLVCPHLNDEYTAEDGTQRCCKSPCARVCYGLCMELLREBAVTS 358
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      300 CPSNNVVE-CAVCYSCAGMLEYD-ENGKRKSCPKCDVCPKCYVDGJGIGSLNFTAVNS 357
OY      359 ANIOEFACCKKIFGSLAFLPESFDDGPASMTAPLOBOLOYFETLEETIGLYISAPDS 418
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      358 TNIRFSMCTKINGIILNRNSFEDBPHKTKGTDPREHMLNLTYYKETGLVLYMMKPN 417
OY      419 LPDISVFQNLQVINGRILLHNGAYS-LTLQGLGISMLGSLRSLRELGSGALILHNHTLCEV 477
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      418 MTSLSVFQNLLETIRGRTFFSGFSEFVVYVQVHNLQWIGLSRLKESVAGVNLKNTLQLRYA 477
OY      478 HTVPMDOLFRRPHQALLTANRPDECEGBELACHOLCARHNCMGPRPOVQVNSQPLRG 537
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      478 NTIMRRFLRSEDDOIEYDART-----EQTGNNNESEDGCM-PGPTMKVCSCLHVRG 529
OY      538 QCEVEECHVLQGLPEEYVNAHCLTCHPCECOPONGSVYTCFEPEDDQCVACAHYKDPFCV 597
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      530 GRCVASCHMLDQEPREAYDGRVCYQHCEGLVQYDLSLTCTYGRFANCSKSAHPPDQOCI 589
OY      598 ARCPGVRPDLSTYMPIMKFPDEEGACORCPITNGHSCVLDLDDKCC 642
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      590 PRCPHGILGDBGDTL-IMKYADRMGCGOCHNCTGOGSGPGLSGC 633

```

RESULT 12  
A42032  
epidermal growth factor receptor - chicken  
C:Species: Gallus gallus (chicken)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 18-Jun-1999  
C:Accession: A42032  
R:Flückinger, T.W.; Maibach, N.J.; Kung, H.J.  
Mol. Cell. Biol. 12, 883-893, 1992  
A:Title: An alternatively processed mRNA from the avian c-erbB gene encodes a soluble  
A:Reference number: A42032, MUID:92123214, PMID:1732751  
A:Accession: A42032  
A:Status: Preliminary  
A:Molecule type: RNA  
A:Residues: 1-527 <PRT>  
A:Cross-references: GB:MT7637; NID:g211737; PIDN:AAA48759.1; PID:g211738  
A:Experimental source: liver  
A:Note: sequence extracted from NCBI backbone (NCBIN:76892, NCBI:P:76893)  
C:Superfamily: epidermal growth factor receptor; protein kinase homology  
C:Keywords: ATP; growth factor receptor

Query Match	22.7%	Score 1155;	DB 2;	Length 527;
Best Local Similarity	44.0%	Pred. No. 1.2e-56;		
Matches 227; Conservative	87;	Mismatches 182;	Indels 20;	Gaps 8;
OY	11	LILALLPGAAST-----OVCHGTMRKLRAPSETHLDMLHYOGGVOGNGLELYL	65	
Db	20	LILLLGRVALCAAVEKKVCGGTNNKLTQLGHVEDHFTSLGRMNCEVVALSNELIYV	79	
OY	66	PTNASTLSFLQDIOEVOGVLLTAHNNQVRPLQRLRIIVGTQTQTFEDNYALAVLNDNDPLNN	125	

```

Db      80 EHNRLDTFLKTIQEVAGVYLALANKVDYPLENIQIRGNVLXNSPALAVLSNYH-MNK 138
Oy      126 TTPVTGASPGGLRELQSLSTEILKGVLIQNRNPOLCYQDTILMKDIFHKNNQALLTLD 185
Db      139 10-----GLRELPMRLSEILNGVAKISNNPKLCNMNDVILMNDIIDTSRK-PLTVLD 189
Oy      186 -TNRSRACHPCSPMKGSRGWGSESDCSLTRPYCAGGCA-RCGSP-PTTQCHOCAG 243
Db      190 FASNISSCPKHPNCTEDHGWAGEONCOTLTKVCAOOCSCGCKVPSDCCHQCCAG 249
Oy      244 CTGPKHSDCLACLHFNHSGICELHCPALVTYNTDTEFSPNPNRGRTYGASCVTACPNY 303
Db      250 CTGPRESQCLACRKRRDADATCKDCPRLVLYNPTTYQMDVNPREGKYSRGATVRCRPNY 309
Oy      304 LSTDVSGCTLVCPPLHNOEVTAEADGTQRCCKSPCARVYCYGLGMHLEVRRAVTSANIQ 363
Db      310 VYTBHSGCVSRGNTDVTYV-EENGVRCKCKDGLCSKVCNGSIGIGELGILSINTNIDS 368
Oy      364 FAGCKKIGSLAFLEPSEFDPASNTAPLOPEQLOVFELEITGYLISAMPDLSPLDLS 423
Db      369 FKNCCKINGDVSILPVALGDAFTTLPDRKLLDYFRTVKEISGFLIQAMPDATILY 428
Oy      424 VFQNLQVIRGILHNGAVSLTLQGLISWGLSLRELGLALHNTLHCFVHTVPMW 483
Db      429 AFENLEIRGRTRKHQGYSLAVVNLKIQSLGLSLKEISDGDIALMKKNLCYADTMNR 488
Oy      484 OLFRRPHALLHTANRPDECVGEGSLACHQACARH 519
Db      489 SLFATQSQKTRKLIQNRKNDK--SKSVCFAPFAKAH 522

```

## RESULT 13

```

A27131
Epidermal growth factor receptor - fruit fly (Drosophila melanogaster) (fragment)
C:Species: Drosophila melanogaster
C:Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 23-May-1997
C:Accession: A27131
R:Schaefer, E.D.; Segal, D.; Glazer, L.; Shilo, B.Z.
Cell 46, 1091-1101, 1986
A:Title: Alternative 5' exons and tissue-specific expression of the Drosophila EGF recep
A:Reference number: A27131; MUID:87002474; PMID:3093080
A:Accession: A27131
A:Molecule type: mRNA
A:Residues: 1-843 <SCH>
C:Genetics:
A:Gene: FlyBase:Bgfr
A:Cross-references: FlyBase:FBgn0003731
C:Superfamily: epidermal growth factor receptor; protein kinase homology
C:Keywords: ATP; growth factor receptor

```

```

Query Match          20.0%; Score 1014.5; DB 2; Length 843;
Best Local Similarity 33.1%; Pred. No. 1.le-48;
Matches 232; Conservative 92; Mismatches 280; Indels 97; Gaps 20;

Oy      24 QVCTSTDMKRLPASPETHLDMRLHLYOGCOVQGNLELTYPT-NASLTFQDIQVOG 82
Db      51 KVCITKRLSLSPSKKEHHYRLNRRTYVDSNKLTLWRPNMNDLSFLDNTREYTG 110
Oy      83 YVLAHNOVOYPLQRLRVKSTOLF-----EDNTALAVLNDGDPDLNNTTPTVGASPGGL 137
Db      111 YLISHVNVKKVFPKLIQIRGLTFLSVEBEKALFY-----TYSKM 154
Oy      138 RELQRLSTEILKGVLIQNRNPOLCYQDTILMKDIFHKNNQALLTLDITNSRACHPCSP 197
Db      155 YTLLEPDLADVLNGVGFHNNYNLCHMRTIQMSELYVNGTDAYVYVDTAPERECPKHE 214
Oy      198 MCKGSRGWGSESDCSLTRPYCAGGCA--RCKGPLPTDCHEQCACTGKSHSDCLAC 255
Db      215 SCTHG-CWGBQPKQAKSKLTCSPQAGRCYGPKPKECHLFCAGCTGTQKQCTAC 273
Oy      256 LHFNHSGICELHCPALVTYNTDTEFSPNPNRGRTYGASCVTACPNYLSLTDVSGCTLV 315
Db      274 KNFEDEAVSKSECPMRKYNPTTYVLETFNPNPGKAYGATCVKECP-GHLLNRNGACVASC 332

```

```

Oy      316 PLHNOEVTAEADGTQRCCKSPCARVYCYGLGMHLEVRRAVTSANIQEPACCKITFGSLA 375
Db      333 PODKKMDKGE-----CVPCNGPCPRTCPGVTLH-----AGNIDSFNCTVLDGNIR 379
Oy      376 FLPESEFDG--DPASNTA-----PLQPOLQVFELEITGYLISAMPDLSPLDSYFON 427
Db      380 ILDTFSGFQDYVANYTNGPRITPLDPERREVFSTYKELTGTGLNLEGHPPORNLSTYRN 439
Oy      428 LQVIRGILHNGAV-SLTLQGLISWGLSLRELGLALHNTLHCFVHTVPMWDLF 486
Db      440 LETIRGQLMESMFALAIIVKSSVLSLEMRNLKQSSSVYIQNRDLCYASNIMPAIQ 499
Oy      487 RNPQALLHTANRPDECVGEGSLACHQACARHCHGPRPTQCVNCSQFLRGQCYEEVR 546
Db      500 KEPEKVVWENLRADLEKNGTISDCNEDGCGAGTDCITCKNFNFNGCTIADCGY 559
Oy      547 LQGLPREVYNAHRCJPCRECPONGSVTCFGEPEADOCVACAHYKDPPEVCARCPGKYR 606
Db      560 ISNAK--FDNRCTKICHPER-----TCNAGADHCOECYHAYDGOHCYSECPK--- 608
Oy      607 DLSTYPIKFPDEEGACQPCPINCTHSCVDDLDDKCPAQRASPLTTSQNEDLGPASPL-- 664
Db      609 -----KYND-RGVCRECHATC-----DGC-----TGPKRTIGIGACTTC 641
Oy      665 -----DSTFRSLLEDDMDGLVDABEYLVPO-QGFPCP 697
Db      642 NLAIINDATYKRCCLKDKCPDY-FWEYVHPQDGSILKP 681

```

## RESULT 14

```

E88257
Protein let-23 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-Aug-2001
C:Accession: E88257
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating bio
A:Reference number: A75000; MUID:99069613; PMID:9851916
A>Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/projects/C-
A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;
A:Accession: E88257
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1323 <STO>
A:Cross-references: GB:chr_II; PIDN:CAA93882.1; PID:g3881523; GSPDB:GNO0020
C:Genetics:
A:Gene: let-23
A:Map position: 2
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom

```

```

Query Match          14.1%; Score 718.5; DB 2; Length 1323;
Best Local Similarity 28.0%; Pred. No. 4.3e-32;
Matches 191; Conservative 107; Mismatches 256; Indels 129; Gaps 28;

Oy      25 VCTGDMKRLPASPETHLDMRLHLYOGCOVQGNLELTYPTN----- 68
Db      39 LCGSTTGINSRYGTGNI-LEDETMYRGCRRYGMELETTWEANIKKRWRESTNSTVDPK 97
Oy      69 -----ASLSPLODQOEVGYVLAHNOVOYPLQRLRVIRVTOFEDNYVALAVDNDGP 122
Db      98 NEDSPKISINFDNLLEIRGSLITIRANIOKISFRLRYTIGDEVFHDN-ALYIKKND- 155
Oy      123 LNTTPTVGASPGGLRELQSLSTEILKGVLIQNRNPOLCYQDTILMKDIFHKNNQALL 181
Db      156 -----VHEVYMRRLRYIRNGSVYIQDNPKMICYIGKIDIMKELLYPD--VQ 199
Oy      182 TLIDTNSRACH-----PSPMKGSRGWGSESDCSLTRPYCAGGCAAC--KQPL 231
Db      200 KYETTNHQHCQYONKSNMAKHESC-NDKCMGSGDNDQRYVRSYCVPSQSCFYSNSTS 258
Oy      232 PTDCHBQCAAGCIGPKHSDCLACLHFNHSGICELHCPALVTYNTDTEFSPNPNRGRTF 291

```

```

Db      259 SYCCCSACAGGCTGCGHKKACIACSKYELDGLCTETCPSKRIENHKGRVLPNDGGRYON 318
OY      292 GASCVTACAPYNYL-SFVGSCTLYC-PLHNOEVTAEEDGTORCEKC-SKPCARVCYGLGME 348
Db      319 GNNVCKECPPELLIENDY--CVRHCSGDHHYDATKD--VRECEKCRSSSCPKITVDG-- 372
OY      349 HLEBRAVTSANTIOEFGACKRIFGSLAFLESPDGDPASTAPLQAPOLQVETLEITG 408
Db      373 HL-----TWETLKNLEGCQIDGHL-TIEHAF-----TYDQLKLVETKIVSE 414
OY      409 YLYISAMPDSLPLDSVQNLQVIRGLIHNGAVSLT-OGIGISWGLRLSRLREISGLAL 467
Db      415 YITIV--QONFYDLKFLKNQIIEGRKLHVRNALAIYQDDLEELSLNLIKIKTAVL 472
OY      468 IHNHTHLCFYHYVPMQDLF---RNPHQALHTANRPDEBCVGBGLACHOLCARHGCMWP 523
Db      473 IMKNHRLCTVSKIDWSITTSKGDKNKPSLAIAENRDSKLCETEQRYVCDKNCKNRGCMWK 532
OY      524 GPTQCVNCSQFLRGQECVEBCRVLQGLPREYVNAHCLPCHPRECOPONGSVTCFGEPAHQ 583
Db      533 EPEDCLECKTKMKSVCYTCVEKCDT-KGFLRNQTSK-CERCSPECE-----TCNGLELD 584
OY      584 CVACAH-----YKDPFCVACRPSGVKPDLSYMPIMKFPDEGACQPCPINCTHSCVD 636
Db      585 CLICRHNKTLVNSDFGNMECVHDC-----PVSHPTQKNVCEKCHPTCY----- 628
OY      637 LDDGCPAEQARASPLTSQNEDELG 659
Db      629 --DNGC-----TGPDNLG 640

RESULT 15
S70712
protein-tyrosine kinase (EC 2.7.1.112) let-23 precursor - Caenorhabditis elegans
N:Alternate names: receptor tyrosine kinase let-23
C:Species: Caenorhabditis elegans
C:Date: 21-Apr-1997 #sequence_revision 09-May-1997 #text_change 11-Jan-2002
C:Accession: S70712; S73101; S13422; T27682
J:Sakai, T.; Koga, M.; Ohshima, Y.
R:Mol. Biol. 256, 548-555, 1996
A:Title: Genomic structure and 5' regulatory regions of the let-23 gene in the nematode
A:Reference number: S70712; M0ID:96177760; PMID:8604137
A:Accession: S70712
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-1374 <SAK>
A:Cross-references: EMBL:D63426
A:Experimental source: strain N2
R:Koga, M.
submitted to the EMBL Data Library, July 1995
A:Reference number: S73101
A:Accession: S73101
A:Molecule type: DNA
A:Residues: 1-50, 'G', 52-1374 <KOG>
A:Cross-references: EMBL:D63466; NID:g1407562; PIDN:BA09729.1; PID:g1407563
A:Experimental source: strain N2
R:Arlian, R.V.; Koga, M.; Mendel, J.E.; Ohshima, Y.; Sternberg, P.W.
Nature 348, 693-699, 1990
A:Title: The let-23 gene necessary for Caenorhabditis elegans vulval induction encodes a
A:Reference number: S13422; M0ID:91080919; PMID:1979659
A:Accession: S13422
A:Molecule type: mRNA
A:Residues: 52-1374 <ARO>
R:Thomas, K.
submitted to the EMBL Data Library, March 1996
A:Reference number: Z20404
A:Accession: T27682
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 52-1374 <WIL>
A:Cross-references: EMBL:Z70038; PIDN:CAA93882.1; GSPDB:GN00020; CESP:ZK1067.1
A:Experimental source: clone ZK1067

```

```

C:Genetics:
A:Gene: let-23; CESP:ZK1067.1
A:Map position: 2
A:Introns: 44/1: 51/1: 90/1: 112/3: 165/3: 217/1: 290/1: 379/1: 418/1: 448/2: 565/1:
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
C:Keywords: ATP; phosphotransferase; transmembrane protein; tyrosine-specific protein
F:1-28/Domain: signal sequence #status predicted <Sig>
F:29-1374/Product: protein-tyrosine kinase let-23 #status predicted <MAT>
F:934-1199/Domain: protein kinase homology <Kin>
F:942-950/Region: protein kinase ATP-binding motif

Query Match      14.1%; Score 718.5; DB 2; Length 1374;
Best Local Similarity 28.0%; Pred. No. 4, 5e-32;
Matches 191; Conservative 107; Mismatches 256; Indels 129; Gaps 28;

OY      25 VCTQTDKMLPLASPEHNDMLRLHYGCQVQVQCNLETLPLPN----- 68
Db      90 LCGSTTGIRYRGNGI-LEDLFTMYRGCRVYNGNLETWIEANETKMKRESTNSTVDPK 148
OY      69 -----ASLSFLDIOEVQGVLIHNOVROVPLQRLIRYGTQLFEDNYALAVLNDGDP 122
Db      149 NEDSPLKSNFNDLEIRGSLIYIRANIQKISPRIRVLYGDEVFHDN-ALYTHKNDK- 206
OY      123 LNNTPVYTGASPGQLRELRLSLTEILKGVLLIQRPOLCY-ODFTIMKDI FHKNOAL 181
Db      207 -----VHEVVRKELRVIRNGSVTLIQDNPKMICYIGDKIDMKELLYDPP--VQ 250
OY      182 TLIDNRSRACH-----PCSPKCKSRGMESEEDCOGLTFYVCAAGGARC--KGPL 231
Db      251 KVEETNSHQHCYONGKSMACCHESC-NDKCMGSDNCCQVYRVSCKSCSQCYNSTS 309
OY      232 PTDCHEQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTVNTDTFESMPNDEGRYTF 291
Db      310 SYECDCACAGCGCTGCGHKKACIACSKYELDGLCTETCPSKRIENHKGRVLPNDGGRYON 369
OY      292 GASCVTACAPYNYL-STVGSCTLYC-PLHNOEVTAEEDGTORCEKC-SKPCARVCYGLGME 348
Db      370 GNNVCKECPPELLIENDY--CVRHCSGDHHYDATKD--VRECEKCRSSSCPKITVDG-- 423
OY      349 HLEBRAVTSANTIOEFGACKRIFGSLAFLESPDGDPASTAPLQAPOLQVETLEITG 408
Db      424 HL-----TWETLKNLEGCQIDGHL-TIEHAF-----TYDQLKLVETKIVSE 465
OY      409 YLYISAMPDSLPLDSVQNLQVIRGLIHNGAVSLT-OGIGISWGLRLSRLREISGLAL 467
Db      466 YITIV--QONFYDLKFLKNQIIEGRKLHVRNALAIYQDDLEELSLNLIKIKTAVL 523
OY      468 IHNHTHLCFYHYVPMQDLF---RNPHQALHTANRPDEBCVGBGLACHOLCARHGCMWP 523
Db      524 IMKNHRLCTVSKIDWSITTSKGDKNKPSLAIAENRDSKLCETEQRYVCDKNCKNRGCMWK 583
OY      524 GPTQCVNCSQFLRGQECVEBCRVLQGLPREYVNAHCLPCHPRECOPONGSVTCFGEPAHQ 583
Db      584 EPEDCLECKTKMKSVCYTCVEKCDT-KGFLRNQTSK-CERCSPECE-----TCNGLELD 635
OY      584 CVACAH-----YKDPFCVACRPSGVKPDLSYMPIMKFPDEGACQPCPINCTHSCVD 636
Db      636 CLICRHNKTLVNSDFGNMECVHDC-----PVSHPTQKNVCEKCHPTCY----- 679
OY      637 LDDGCPAEQARASPLTSQNEDELG 659
Db      680 --DNGC-----TGPDNLG 691

```

Search completed: January 13, 2003, 14:49:14  
 Job time : 33.5383 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2003, 14:44:35 ; Search time 42.2593 Seconds

(without alignments)  
4480.838 Million cell updates/sec

Title: US-09-854-356-6

Perfect score: 5078  
Sequence: 1 METALCRNGLLALLPPGA.....TFKCTPAENPEYLGLDVPY 919

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21:\*  
1: sp.archaea:\*  
2: sp.bacteria:\*  
3: sp.fungi:\*  
4: sp.human:\*  
5: sp.invertebrate:\*  
6: sp.mammal:\*  
7: sp.mhc:\*  
8: sp.organelle:\*  
9: sp.phage:\*  
10: sp.plant:\*  
11: sp.podent:\*  
12: sp.virus:\*  
13: sp.vertebrate:\*  
14: sp.unclassified:\*  
15: sp.virus:\*  
16: sp.bacteriap:\*  
17: sp.archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4438	87.4	1259	6 O18735	O18735 canis fam1
2	1874.5	36.9	419	4 Q9UK79	Q9UK79 homo sapien
3	1631	32.1	1209	11 Q9GX70	Q9GX70 rattus norv
4	1609	31.7	1210	11 Q9EP98	Q9EP98 mus musculu
5	1533.5	30.2	655	11 Q9WVF5	Q9WVF5 mus musculu
6	1517.5	29.9	643	11 Q9ERV6	Q9ERV6 mus musculu
7	1357	26.7	1328	13 P79754	P79754 fugu rubrip
8	1353	26.6	1165	13 Q9YH40	Q9YH40 xiphophoru
9	1252.5	24.7	1137	13 Q9W6F6	Q9W6F6 gallus gall
10	1217	24.0	367	11 O8R2X1	O8R2X1 mus musculu
11	1175.5	23.1	412	4 Q8WYV0	Q8WYV0 homo sapien
12	1155	22.7	537	13 Q90836	Q90836 gallus gall
13	1032	20.3	1433	5 Q9B1H9	Q9B1H9 anopheles g
14	1028.5	20.3	478	11 Q9ESD0	Q9ESD0 rattus norv
15	971.5	19.1	599	13 Q9PSH2	Q9PSH2 gallus gall
16	901	17.7	165	4 Q14256	Q14256 homo sapien

17	887	17.5	176	11 Q923V5	Q923V5 rattus norv
18	734	14.5	331	4 Q9BDU7	Q9BDU7 homo sapien
19	723	14.2	149	6 Q9BG66	Q9BG66 oryctolagus
20	656	12.9	1368	5 Q23821	Q23821 caenorhabdi
21	621	12.2	1717	5 Q26566	Q26566 schistosoma
22	620.5	12.2	1193	5 Q9Y1X8	Q9Y1X8 ephydaria f
23	409.5	8.1	1362	13 Q9PVZ4	Q9PVZ4 xenopus lae
24	404.5	8.0	150	6 Q9BG64	Q9BG64 oryctolagus
25	395	7.8	151	6 Q9BG65	Q9BG65 oryctolagus
26	366.5	7.2	1368	13 Q8UW85	Q8UW85 parailichthy
27	366.5	7.2	1472	5 Q9U5A8	Q9U5A8 bombyx mori
28	363.5	7.2	366	5 Q26569	Q26569 schistosoma
29	362	7.1	1369	13 Q8UW86	Q8UW86 parailichthy
30	359	7.1	1358	13 Q73798	Q73798 xenopus lae
31	350.5	6.9	136	11 Q99PK1	Q99PK1 mus musculu
32	345.5	6.8	1418	13 Q9J457	Q9J457 scophthalmu
33	340.5	6.7	1671	5 Q9NJV5	Q9NJV5 biophthalari
34	339.5	6.7	410	11 Q63720	Q63720 rattus norv
35	331	6.5	334	5 Q26567	Q26567 schistosoma
36	331	6.5	342	5 Q26568	Q26568 schistosoma
37	328	6.5	149	11 Q60494	Q60494 cavia sp. e
38	326.5	6.4	1245	13 Q9YGH8	Q9YGH8 scophthalmu
39	325	6.4	1418	13 Q8UW83	Q8UW83 parailichthy
40	322	6.3	469	11 Q63721	Q63721 rattus norv
41	316	6.2	1371	11 Q9QVW4	Q9QVW4 rattus sp.
42	315	6.2	1412	13 Q8UW84	Q8UW84 parailichthy
43	307	6.0	2144	5 Q9VDP4	Q9VDP4 drosophila
44	302.5	6.0	1749	5 Q8TOW6	Q8TOW6 echinococcu
45	296	5.8	946	5 Q9VJ04	Q9VJ04 drosophila

#### ALIGNMENTS

RESULT 1  
ID O18735 PRELIMINARY; PRT: 1259 AA.  
AC O18735:  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE ErB-2.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_Taxid=9615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Yokota H.;  
RT "cDNA cloning of erB-2 from canine mammary gland.";  
RL Submitted (Oct-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB008451; BAA23127.1; -.  
DR HSSP: P11362; 1FGK.  
DR InterPro: IPR002048; EF-hand.  
DR InterPro: IPR000494; EGFR\_Ldomain.  
DR InterPro: IPR000719; Euk\_pkinase.  
DR InterPro: IPR002174; Furin-like.  
DR InterPro: IPR001245; Tyr\_pkinase.  
DR InterPro: IPR004019; YLP\_motif.  
DR Pfam: PF00757; Furin-like; 1.  
DR Pfam: PF00069; kinase; 1.  
DR Pfam: PF01030; Recep\_Ldomain; 2.  
DR Pfam: PF02757; YLP; 2.  
DR ProDom: PD000001; Euk\_pkinase; 1.  
DR SMART: SM00261; FU; 3.  
DR SMART: SM00219; Tyrc; 1.  
DR PROSITE: PS00018; EF\_HAND; UNKNOWN; 1.  
DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE: PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
KW ATP-binding; Transferase; Tyrosine-protein kinase.  
SQ SEQUENCE 1259 AA; 137989 MW; E37364D9C4ACDA6 CRC64;

Query Match	87.4%	Score 4438	DB 6	Length 1259
Best Local Similarity	66.4%	Pred. No. 0		
Matches 837	Conservative 31	Mismatches 50	Indels 342	Gaps 3
QY 1	MELALRCRGLLTLALPPCAASTQVCTGDMKLRIPASPTNIDMLRHLYGSCGVVQGNL	60		
Db 1	MELAMCKRGLLTLALPPSAAGTQVCTGDMKLRIPASPTNIDMLRHLYGSCGVVQGNL	60		
QY 61	ELTYLPTNASLSFLDIOEVQGVYLIANNOVAPQLRILRYGTOLFEDNALAVLNDG	120		
Db 61	ELTYLPANASLSFLDIOEVQGVYLIANSQVAPQLRILRYGTOLFEDNALAVLNDG	120		
QY 121	DPLNNTFVTGASPGGLKRELQRLSLTEILKGGVLLIQRPNQLCYDTILMKDIFHKNNQIA	180		
Db 121	DPLEGIGPARGAAGQRELQRLSLTEILKGGVLLIQRSFOLCHODPIIMKDVHKKNNQIA	180		
QY 181	LTLIDTNSRACHPSMCKSRGCGESSEDCQSILTRYCAGGACRCKGRPLTDCHEQC	240		
Db 181	LTLIDTNSRACHPSMCKSRGCGESSEDCQSILTRYCAGGACRCKGRPLTDCHEQC	240		
QY 241	AAGCTGPKHSDCLACILHFNHSGICELHCPALTYMTDFEPMNPBEGRTFGASCVTACP	300		
Db 241	AAGCTGPKHSDCLACILHFNHSGICELHCPALTYMTDFEPMNPBEGRTFGASCVTSCP	300		
QY 301	YNYLSTDVGSCSTLVCPRLHNOEVTABDGTORCEKSKPCARVCGYGLGMEHLREAVTASN	360		
Db 301	YNYLSTDVGSCSTLVCPRLHNOEVTABDGTORCEKSKPCARVCGYGLGMEHLREAVTASN	360		
QY 361	IOEFAGCKKIFESLAFPERSPDGDASNTAPLQRPOLRFELTEETGLYLSAMPDSLAP	420		
Db 361	IOEFAGCKKIFESLAFPERSPDGDASNTAPLQRPOLRFELTEETGLYLSAMPDSLAP	420		
QY 421	DLSVFNQNLQVILGRILHNGAVSLTLQGLGISWLGRLSELTSGLALIHNNHLFCVHVY	480		
Db 421	NLSVFNQNLRVILGRVLDHNGAVSLTLQGLGISWLGRLSELTSGLALIHNNHLFCVHVY	480		
QY 481	PMDQLFRNPQHALLTANRPDECEVGEGLACHQLCARGHCWGPRTQVCNCSQFLRGDEC	540		
Db 481	PMDQLFRNPQHALLTANRPDECEVGEGLACHQYR-CANGHCMWGPRTQVCNCSQFLRGDEC	539		
QY 541	VEECVNLQGLPREYVYANHCLPCHEQCQPNQNSVYCFGEADQCYACAHYKRRPCVVAR	600		
Db 540	VEECVNLQGLPREYVKDKRCLPCHEQCQPNQNSVYCFGEADQCYACAHYKRRPCVVARC	599		
QY 601	PSGVKPRDLSYMPIMKFPDEEGACQCPILNCTHSYCDLDDKCGAPARASPLTS-----	653		
Db 600	PSGVKPRDLSYMPIMKFPDEEGTCQCPILNCTHSYCDLDDKCGAPARASPLTSITAAVVG	659		
QY 654	-----	653		
Db 660	ILLAVVGLVLGILIKRRRQKIRKYTMRLQETELVEPLTPSGAMPNQAQMRILKETEL	719		
QY 654	-----	653		
Db 720	RKVVLGSGAGSTYKGIWIPDGENVKIPVALIKVIRENTSPRANKELIDEAYVMAGVSP	779		
QY 654	-----	653		
Db 780	YVSRLGLGICLSTVQLVQLMYGGLDHNHRRRLQSODLLMNCYQIAKGMSTLEDEVY	839		
QY 654	-----	653		
Db 840	LVRHDLARNVLVKSPNHVKITDFGLARLLDIDETEHADGKVPDKMMALLESIPRRPT	899		
QY 654	-----	653		
Db 900	HOSDWSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGERLRQRPICITIDVYIMVYKWM	959		
QY 654	-----	653		
Db 960	IDSECRPRFRELVAEFSSMADPQRFVYIQNEDLSPASPLDSTFTRSLLEDDMDGLDVA	1019		
QY 685	EELYLVPOGFCPPAPAGAGVHHRRHSSSTRSGGDLTLGLFESEEARSPPLAPSEG	744		

[illegible]

[illegible]

D	b	125	-----YGNKTKGABELPMRNLQELLIGAVRFNSNNPLICNMETIQMRDIY-QDVPLSN	175
Q	y	182	TLIDTNRN-RACHPCSPCKSGSRGWGESSDQSLRTYACAGGA-RCKGRPLPTDCHQ	239
D	b	176	MSMVOHRLTGCPKCDSPCSGSGSGWGEENCOIKTIIICAQDCSRGRGRSPDCHNQ	235
Q	y	240	CAAGCTGPKHSDCLACJHFHNSGICELHCALTYMTDPFESMNPNEGRTFGASGVAC	299
D	b	236	CAAGCTGRESDCLVCRFDEATCKDTCPLMLYPTTYQMDVNPBGKTSFGATCVKCC	295
Q	y	300	PYNTLSTDVGSCTVCPFLHQEYTAEDGTQRCCKSGPCARVOCYGGIMHLEREVAVTA	359
D	b	296	PRNVVVDHSGCVACACPDYEV-EBDGVSKCKCQCPCKRKYONGIOGFEDTISNAT	354
Q	y	360	NIOFAGCKKIIFGSLAFIPESFDGDPASNTAPLOBOLOYFETLEETIGLYISAMPDL	419
D	b	355	NIKHFKYCTAISGDLHILPAFKGDSFTPTPDPRELEITLKVKEITGFTLLIQAPENW	414
Q	y	420	PDLSVPQNLQYIRIRILHNGAYSTLQGLISMLGRLSRELSGLALHNTHLCFVNT	479
D	b	415	TDLAFEFLEIRTRTKOHQOFSLAVGNLITSIGLSKEISDGDVITISGNRNLCYAT	474
Q	y	480	VPMQLFRNPQOALLHTANPEDECEVGGIACHQOLARGHCMPGPTQCVNSQFLRQ	539
D	b	475	INMKKLEGTPNQKTKIMANNAEKDKRATNVNCPILSSBECMGPBETDCVQCNNSRGE	534
Q	y	540	CVEBCRYLQGLPREYVNAHRCPCHEPCQDPONGSVTCFGEADQCVAAHYKDPRECVAR	599
D	b	535	CVDKNILEGPRFEFVENSECICQHEPCLPQTMNITCTGKRPNCIKCAHYVDGPHCVT	594
Q	y	600	CPSGVKPDLSTMPYMKRPDEGACQPCPINCITSQVDDDKG-	642
D	b	595	CPSGIMEGNNTL-WKFPADANNVCHLCHANCYTGCGPGLKGQDQEGPKIPSTIGTVG	653
Q	y	643	-----PAQORASPLTQSN-	656
D	b	654	GLFTVVAALGIGLFEMRRQDLYKRIIRLQRELYEBLTPSGEAPNOAHRIILKETEF	713
Q	y	657	-----	656
D	b	714	KKIVLGSAGFVYKGLMIPBEGKVYIPAIKELREATSPRANKRIIDEAVVMSVDNP	773
Q	y	657	-----	656
D	b	774	HVCRLGICLTSTVOLITQMLPQGLDLYVREHKDNGSQYLLMNCVQIAKMNLYEDRR	833
Q	y	657	-----DLG-----	659
D	b	834	LVRDLARNVLYKTPQHVKITPQITFGIAKLGAEBEKEYHAEGGKVPKIMNALESILHRTY	893
Q	y	660	-----PA-----	661
D	b	894	HQSDVMSYGVYWMELMTFGSKPYDGIPIASISILEKGERLQRPICITIDVYMINVKCM	953
Q	y	662	-----SPLDSTFYSLLEDDMGDLVD	683
D	b	954	IDADSRPKFRELILFESKMARDPQRYLQGDERMHLPSPTDSNFYALMEBEDMEDYVD	1013
Q	y	684	AELTYVQOQGFPCDPRPAGAGVNHHRHSSSTRSGGDLITGLPESSEAPARSPALASE	743
D	b	1014	ADELTYVQOQGF-----NSPT-----SRTPLSL	1039
Q	y	744	GAGSDVDDGDLGMCAGKGLSLPTHPSPFORXSEBPTVPLPSET--DGVVAPLCSPOP	801
D	b	1040	SANS-----SSTYACINRNSCKVAKEDAFLOKRTSSPPTSVLEEDNIDDTFL-----FYP	1088
Q	y	802	EYVNPQVPRQPPSPREGPLPAARPAAGTL-----ERPTLSPGKNVYKVDFAFGAVE	856
D	b	1090	EYINQ-----SVF-KRPAQSVONPYHNOPLHAPGRDLHYON--PSHNAVS	1133
Q	y	857	NPEYL-TPQGGAPQPHRPAPSPADNLXYNDQ-----DP-----PENGAPPS	899
D	b	1134	NPEYLNLAQ-----PPLCSSGSSSALWLOKSGHOMSLDNPDYQODFPKAKPNG	1184





DB 1044 SN-----NSTVACINRNGSCRVKEDAFLOKRYSDPTGAVTEDNIDAF-----PVEPYV 1093  
 QY 805 NQDPVAPPPSPREGLPARAPAGATL-----ERPRTLSPGKNVYKDYAFGCAVENPE 859  
 DB 1094 NQ-----SVP-KRPAQSVQNPVYHNPQLHPAPGRRLATON--PHSNANGNPE 1137  
 QY 860 YL-TPGGGAAPQHPHPAPSPAFDNLXYWDQ-----DP-----PERGAPSTTK 902  
 DB 1138 YLNTAQ-----PTCLSSGSPNPAIMIQKSHQMSLDNDYQDDFFPKETKPKMGIRK 1188  
 QY 903 GTPPAENPEYGLDVP 918  
 DB 1189 G-PTAENAEYLRVAP 1203  
 RESULT 5  
 Q9WVF5 PRELIMINARY; PRT; 655 AA.  
 AC Q9WVF5;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Epidermal growth factor receptor (Epidermal growth factor receptor isoform 3).  
 GN EGFR.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-C57BL/6J; TISSUE-LIVER;  
 RA Reiter J.L., Threagill D.W., Danielson A.J., Schell C., Lampand A.L., Balasubramaniam S., Crossley T.O., Magnuson T.R., Maille N.J.;  
 RT "Alternative transcripts from the Human and Mouse EGFR Genes Encode Carboxy-Terminal Truncated Receptors."  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 RN [12]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-C3H/101, 129/SVJ, AND 129/SVEV7AC;  
 RA Reiter J.L., Threagill D.W., Eley G.D., Strunk K.E., Danielson A.J., Sinclair C.S., Pearsall R.S., Green P.J., Yee D., Lampand A.L., Balasubramaniam S., Crossley T.O., Magnuson T.R., James C.D., Maille N.J.;  
 RT "Comparative genomic sequence analysis and isolation of human and mouse alternative Egr transcripts encoding truncated receptor isoforms."  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 RN [13]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-C57BL/6J; TISSUE-LIVER;  
 RA MEDLINE-21085660, PubMed-11217851;  
 RA Kawai J., Shitagawa A., Shibata K., Yoshino M., Itoh M., Ishi Y., Aizawa T., Hara A., Fukunishi Y., Konno H., Aachi J., Fukuda S., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka T., Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischman W., Gaasterland T., Gissi C., King B., Kochia H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schiml L.M., Staubl F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boileau D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Schoenwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmink D., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuki S., Hayashizaki Y.;  
 RA "Functional annotation of a full-length mouse cDNA collection."  
 RT Nature 409:685-690(2001).

DR EMBL: AF124513; AAD4149.1; -  
 DR EMBL: AF275366; AAG28047.1; -  
 DR EMBL: AF275364; AAG28047.1; JOINED.  
 DR EMBL: AF275365; AAG28047.1; JOINED.  
 DR EMBL: AK004944; BAB23688.1; -  
 DR EMBL: AK004883; BAB23641.1; -  
 DR EMBL: AK004911; BAB23662.1; -  
 DR MGI:95294; Egr.  
 DR Interpro: IPR000494; EGFR\_L\_domain.  
 DR Interpro: IPR002174; Furin\_Like.  
 DR Pfam: PF00757; Furin-Like; 1.  
 DR Pfam: PF01030; Recep\_L\_domain; 2.  
 DR SMART: SM00261; FU; 3.  
 KM Receptor.  
 SQ SEQUENCE 655 AA; 72906 MW; 6B34063B1BC928CB CRC64;  
 Query Match 30.2%; Score 1533.5; DB 11; Length 655;  
 Best Local Similarity 44.4%; Pred. No. 5,8e-111;  
 Matches 284; Conservative 99; Mismatches 233; Indels 23; Gaps 7;  
 QY 11 LIALLPQAA--STVCGTGMKRLPASPETHLDMHLHYOGQYVGNLELYLPTN 68  
 DB 14 LTLALCAAGALEERKVCQGTNRNLQIGTFEDHFLSLQRMYNCEVVLGNLEITYVQNR 73  
 QY 69 ASLSFQDIOEVQYVLIHNOYRQVPLQRLIRVGRTOLEFQNVALAVDNGDPLNNTTP 128  
 DB 74 YLSLFTKTIQEVAGYVLIHLNLYVERLPLENLQIRGNALYENTYALALISN----- 124  
 QY 129 VYGASPGRLRELQRLSLRELKGVLIQRPOLCYODTILMDI----PHKNNQALTLI 184  
 DB 125 -YGTNRGTRELPMNLQILIGAVFNSNPILCNMDTIQWMDIYQVFMNSMDL--- 180  
 QY 185 DNRKSRACHPGSPMKSGKSGWGESSEDCSLTRIVYAGCA-RCKGPLTDCCHQCAAG 243  
 DB 181 -QSHPSCKPCPCSPGNSCWGSENCOKLTKITCAQCSHRCRSPSDCHQCAAG 239  
 QY 244 CTGPKRSDDLACLFHNSGICELHCPALVTYNTDPESMPNDEGRTPGASCVTACPYV 303  
 DB 240 CTGPRSDCLVQCKRQDEATCKDTCPPMLYNTPTTQMDYVNEBGKTSFATCKKCPRY 299  
 QY 304 LSTVGSCTLVCPNLHNOEYTAEDGTQRCESKPCARCYGCMGHELRVAVTSANIOE 363  
 DB 300 VYDHSQVRAAGPDYEV-EDGIRKCKCGPKVCNGIGEGFNTLSINATNIK 358  
 QY 364 PACCKRTFGLAFPEPSFGDPAANTAPLQPEQLOFETLEITGLYLSANPDSLPDIS 423  
 DB 359 FYCTAISGDLHLPLPAFKGDSFTTRPLDPRELEITLTKVKEITGLLIQAMPDMWTDLH 418  
 QY 424 VFQNLQVINGRLIHNGAVSLTLOGISWLSGRLSRLGSLALHNNHNLQFVHTVPM 483  
 DB 419 AFEENLEITGRKQHGQSLAVAGLITSLGRSLKEISDGVILISGNRLCYANTINK 478  
 QY 484 QLEFRNQALLHTANRPDECEVGEGLAQHQLARGHGWPGTQCNSQSFQGEVVE 543  
 DB 479 KLFGRNQKTKIMNNAEKDKAVNHYCNPLSSSECGWPERDVCVQGNVSRGRCVEYK 538  
 QY 544 CRYLQGLPREYVNAHCLPCHPGCPQNSVTCFGEADQVACAHYKDPVCVARGSPG 603  
 DB 539 CILIEEPREFEYNSCLOCHCECLPQAMNITCTGGRPNCIOCAHYIDGPRCVTQCPAG 598  
 QY 604 VKPDLSTYPMIKFPEEGACQPCPTNCTHSCVDLDKGC 642  
 DB 599 IMGENNTL-VMKYADANNVCHLCHANCYGCAGPGIQC 636  
 RESULT 6  
 Q9ERV6 PRELIMINARY; PRT; 643 AA.  
 AC Q9ERV6;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Epidermal growth factor receptor isoform 2.



```

GN EGFR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C3H/101, 129/SVJ, AND 129/SVEV7AC;
RA Reiter J.L., Threadgill D.W., Elen G.D., Strunk K.E., Danielson A.J.,
RA Sinclair C.S., Pearsall R.S., Green P.J., Yee D., Lampard A.L.,
RA Balasubramanian S., Crossley T.O., Magnuson T.R., James C.D.,
RA Maibhe N.J.;
RT "Comparative genomic sequence analysis and isolation of human and
RT mouse alternative Egfr transcripts encoding truncated receptor
RT isoforms.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF215366; AAC28046.1; JOINED.
DR EMBL; AF215364; AAC28046.1; JOINED.
DR EMBL; AF215365; AAC28046.1; JOINED.
DR MGD; MGI:95294; Egfr.
DR InterPro; IPR000345; CytC_heme_bind.
DR InterPro; IPR000494; EGFR_L_domain.
DR InterPro; IPR002174; Furin-like.
DR Pfam; PF00757; Furin-like.
DR Pfam; PF01030; Recep_L-like; 1.
DR SMART; SM00261; FU; 4.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW Receptor.
SO SEQUENCE. 643 AA; 71476 MW; DEF22002C84911B1 CRC64;

Query Match 29.9%; Score 1517.5; DB 11; Length 643;
Best Local Similarity 44.7%; Pred. No. 1e-109;
Matches 281; Conservative 98; Mismatches 227; Indels 23; Gaps 7;

QY 11 LLLALLPAGAA--STQVCTGTDMLRLPASBETHLDMRLHLYQCGOVQGNLETLPLN 68
DB 14 LLLALCAAGALAEKKVGGSTNRLTQGTGDFHLSLQRMVNCVAVGNLEITLYORN 73
QY 69 ASSLFLDIOEYGVYLLAHNOVROVPLORLRIYRGTLQDFEDNVLAVIDNDPLINNTTP 128
DB 74 YDLSLTKTIOEYAGVYLLALMTVERIPLENQIINGNALYETATLALISN----- 124
QY 129 VTGASPGRLQSLRLSLLEILKGVLLIQNPOLCYQDTILMKDI---FHKNNQALTLI 184
DB 125 -YGTNRTGLRELPMRNLIQELIIGAVFSNNPILCMQDTIQRMDIYQNVFMSMSMDL--- 180
QY 185 DTNRARARPCSPMKSCRCWGESSEDCQSLTRYCAGGCA-RCKGRLPTDCHQCGAAG 243
DB 181 -QSHSPCKPCDPSCPNCSGCGGEGENCOKLTKIICAOCCSHRCGRSPSDCHNQCAAG 239
QY 244 CTGPKHSQCLACLNHNHSGICELHCPALVTYNTDTFESMPNDEGRYTFGASCVTACPNY 303
DB 240 CTGPRESCLVQKQODATCKDTCRPLMLNPTTYQMDVNPDEGRYSFGATGVKKCPKNY 299
QY 304 LSTDVGSCTLVCPRLNHOEYTAEDGTQRCCKSPCARCYGIGMHLREAVRYTSANIOE 363
DB 300 VVTDHSGSVRCAPDYEV-EEEDGIRKCKKCDGRCRKYCNGIGIEFDTLSINATNFKH 358
QY 364 FAGCKKIFGSLAFLPESFDGDPASNTAPLQPOLOVFTLEBTIGYLYISAMPDLS 423
DB 359 FKCYTAISGDALILPVAKGDSFTETPLDRELEILTKVEITGFLIQAAMPDWTLDH 418
QY 424 VFQNLQVIRGLINNGAVSLTLOGIGISMLGRSLRELGSALTNHNTHICFVNTVMD 483
DB 419 AFENLEILRGTRKQHGQSLAVNGLNTSLGRSLKETSDGVITISGRNNLCYAFITWK 478
QY 484 QLFNRPHQALHTANRPDECVGEGLACQLCARGHCWGPRTQCVNCSQFLQGECEVE 543
DB 479 KLFGRPNKTKIMNRRAKCDKAVNHYCNPLCSSSGCGRPREDVCSQNNSRGCEYK 538
QY 544 CRVTLQGLREYVNAHRLCPHREPCQPNQSVYTCFGRPADQCVACAHYKDPPECVACPSG 603
DB 539 CNILEGPERFEVNSSECIOCHPECLPQAMNITTCGRGPDNCIOCAHYIDGPHCVTCDAG 598

```

```

QY 604 VKPDLSTYMPWKPEDEGACQPCPINTH 632
DB 599 IMGENTTL-VMKYADANNVCHLCHANCY 626

RESULT 7
ID P79754 PRELIMINARY; PRT; 1328 AA.
AC P79754.
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Erib3.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RL MEDLINE=99177347; PubMed=10077531;
RA Gellner K., Brenner S.;
RT "Analysis of 148 kb of genomic DNA around the wnt1 locus of Fugu
RT rubripes.";
RL Genome Res. 9:251-258(1999).
DR EMBL; AF056116; AAC34391.1; -.
DR HSSP; P11362; IFEK.
DR InterPro; IPR000494; EGFR_L_domain.
DR InterPro; IPR000719; Euk_kinase.
DR InterPro; IPR002174; Furin-like.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF00069; Kinase; 1.
DR Pfam; PF01030; Recep_L_domain; 2.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00261; FU; 3.
DR SMART; SM00219; TyrcKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Transferase.
SO SEQUENCE. 1328 AA; 148613 MW; A333039258B647E9 CRC64;

Query Match 26.7%; Score 1357; DB 13; Length 1328;
Best Local Similarity 37.7%; Pred. No. 8.5e-97;
Matches 317; Conservative 101; Mismatches 308; Indels 114; Gaps 26;

QY 9 WGLIALLP--GAASDQ---VCTGTDMLRLPASBETHLDMRLHLYQCGOVQGNLEL 62
DB 4 WRLILMCVASHLRASSTQIEAVCPGTONGLSSTGSGENQYNLNKDRYKGEIIMGNLEI 63
QY 63 TYLPTNASTSLDIOEYGVYLLAHNOVROVPLORLRIYRGTLQDFEDNVLAVIDNDP 122
DB 64 TOIESNMPFSLKTIKRETYGYLLAHNFHOEIPILQQLVIRGNSLYERRFALSVELN--- 120
QY 123 LNTPTVAGSPGRLREIOLNLSIEILKGVLLIQNPOLCYQDTILMKDI---FHKNNQALTL 182
DB 121 ---YKDG--PSGILNQLGMLNLTETILDGVOIINNKILRYGPPVYWDIT--RNNDAITE 173
QY 183 LIDTNRARARPCSPMKSCRCWGESSEDCQSLTRYCAGGCA-ARCKRLPTDCHQCGA 241
DB 174 IQFNREBVCVH---KSC-GNVCWGKQKQOQILITVCAPOCNDRCFTSPRDCCHIECA 229
QY 242 AGCTGPKHSQCLACLNHNHSGICELHCPALVTYNTDTFESMPNDEGRYTFGASCVTACPNY 301
DB 230 AGCKRLPTDCEACFLFNDGACVYQCPQTLTYNQRTQMETNPAKYQYSGISCVSQCP 289
QY 302 NYLSTDVGSCTLVCPRLNHOEYTAEDGTQRCCKSPCARCYGIGMHLREAVRYTSAN 360
DB 290 HFV-VDGSVCVYCPDPKMEV--ERGSQRCQELSSGLCPKVECGAG--QROTVSSN 343
QY 361 IQFAGCKKIFGSLAFLPESFDGDPASNTAPLQPOLOVFTLEBTIGYLYISAMPDLS 420

```

```
Db 344 IDSEFNCTKIQGSLHFLVYGLIGDDDFKNVPELDAKKEVFRYTRITDLINISWKELN 403
Qy 421 DLSVPONLOVIRKRIILHNGAYSLTLOGLSIWLGLSLRELSGLALLIHHNHLCTVHY 480
Db 404 DLSVFSSTLTIOGRSLFKRFRSLVMKRIPTLTSLGLRSLREISDGSVYISQNALCYHHYV 463
Qy 481 PMQLFRNRP-HQALLHTANRPEDECEGLACHQLCARGHGMGPGLTQCNCQOPLRGOE 539
Db 464 NMQLRGRSVRANSLNSNRPMAECVADGVCDPLCSDSCWGPDPQCLSRNYSRHGT 523
Qy 540 CVEECVILQGLPREVYNARH-CLPCHPECOPONGSVTCFPEADQCACAHYKDPFECVA 598
Db 524 CVAGCHFNISGIRPREFAGLNGVCVACHPECKPQTGKASCTGPGADEMACIKFNDGYCMS 583
Qy 599 RCPGVKPDLSYPIWKFPDEBAGACOPPCINCTHSCVDLDDKCPAQRASPLTSGNEDL 658
Db 584 SCPAGVN-DGEKGLIFKFPNRREGHCPCQONCTQCGSPGLNDCLBAARLTISSQITGI 642
Qy 659 GPASPLDSIFRYRSLLEDMDGDLVDAEYLVPOGCFPCDPAPGAGMHNHR-----HRSS 714
Db 643 ALGVPRGLIFCLVL-----FFL-----GMLYHRLALIRKR 673
Qy 715 STRSGGDLTLGLEPSEEBAPRSP-----LAPSE-----GAGSDVFPDGLMG- 757
Db 674 AMNR-----YLESGSEFPLRGEGKGTKHARILKPDILKIRPLGSGVF-GTYSKGF 725
Qy 758 -----AAKGLQSLPTHDPSPLORYSE--DPTVPLPSETDGY-----VAPLTC 797
Db 726 WPEGETVIRPAIKRIQ-----DSSGRQTFEITDHLTMSGLDHPYIVRLIGICPGTC 780
```

```
RESULT 8
09TH40 PRELIMINARY: PRT: 1165 AA.
AC 09YH40:
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Receptor tyrosine kinase, proto-oncogene.
GN XMRK.
OS Xiphophorus xiphidium.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Neuteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha;
OC Cyprinodontiformes; Poeciliidae; Xiphophorus.
OX NCBI_Taxid=8086;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RIO PURIFICATION:
RX MEDLINE-98241172: PubMed-9582016;
RA Dimitrijevic N., Winkler C., Wellbrock C., Gomez A., Duschl J.,
RT Altschmid J., Scharl M.;
RT "Activation of the Xmrk proto-oncogene of xiphophorus by
RT overexpression and mutational alterations.";
RT Oncogene 16:1681-1690(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-RIO PURIFICATION:
RA Scharl M.;
RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: U53471; AADI0500.2; -.
DR HSSP: P11362; IFGK.
DR InterPro: IPR000345; Cyrc_heme_bind.
DR InterPro: IPR000494; EGFR_L_domain.
DR InterPro: IPR000719; Euk_Pkinase.
DR InterPro: IPR002174; Furin-like.
DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR001245; Tyr_kinase.
DR Pfam: PF00757; Furin-like; 1.
DR Pfam: PF00069; pkinase; 1.
DR Pfam: PF01030; Recep_L_domain; 2.
DR PRINTS: PR00109; TYRKINASE.
```

```
DR Prodom: PD000001; Euk_pkinase; 1.
DR SMART: SM00261; RU: 3-.
DR SMART: SM00219; Tyrc; 1.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 2.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
KW ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.
SQ SEQUENCE 1165 AA; 129614 MW; 7FEE38D871A7AE CRC64;

Query Match 26.6% Score 1353; DB 13; Length 1165;
Best local similarity 42.1% Pred. No. 1.5e-96;
Matches 273; Conservative 91; Mismatches 258; Indels 27; Gaps 12;

Qy 1 MEIALACRWGLLALLPFG-AAST---OVCTGTDMKRLPASPETHLMDLRLLYOGCOV 55
Db 4 LELLLEL-----LILLLLISGRCCSTDPDRKVCQSTNQMTM---LDNHLYLKKMKMYSCNV 56
Qy 56 VQGNLELTLYLPTNASSLPIQDIOEVGYLLIANOVROYPLQRLRYRGTOLEFEDYKALA 115
Db 57 VLENLEITYOEQDLSFLQSIQIEVGVYLIAMNEVSTIPVLVRLIRQONLYEGNFTLL 116
Qy 116 VLDNGDPLNNTPTVTCASPGSLRELQRLSTLETILAKGVYLIRNPOLCYQDITLMKDIPIK 175
Db 117 VMSNTYCK-HPSSP--DYVOGLKQLQDLSNITELISGVKYSNHPILCNVETIMMDIYDK 173
Qy 176 NNQALTLIDITNRSRACHPCSPCKSGRCWGESSEDCQSILTRTVCAAGC-ARCKGRLPND 234
Db 174 TSNPTNLLIPHAERQCCQKCDPCVNGSCWAPRGHCQFTKLKLAEOCNRRCRGPBPD 233
Qy 235 CCHQCAAGCTGPKHSDCLACHFNHSGICELHCPRLVYVYNDTFESMNPBERITFGAS 294
Db 234 CCNEHCAGCTGGRATDCLACRFNDGTCCKDPCPKLYDVISQVYDNPNIKYTFGA 293
Qy 295 CVATCPYNTLSTPDVSGCTLVCPILNQEVTAEDTQCEKSKRCPCAVCYGLMEHLREVR 354
Db 294 CYKEGCSNIVYIE-GACVASCASGMLEVP-ENKRSCKRCQDVCPCVCGIGIGISLNTI 351
Qy 355 AVTSANIQEFACCKIFGSLAPLPESEDDDPASNTAPLOPELOVETLEETGYLYISA 414
Db 352 AVNSTIGSFSNCTKINGDIILNRNSFEEDDPHYKIGPMDPEHMLNTYKELTYGLVIMW 411
Qy 415 WPDLSPLDSVPONLOVIRKRIILHNGAYS-LTLOGLSIWLGLSLRELSGLALLIHHNHN 473
Db 412 WPENMSTLSVFOLELIRRTFTSRGFSFVVQVSHLQWLGLSSLKEVASAGNVILKNTPO 471
Qy 474 LCFVHTVPMDQLFRNPHQALLHTANRPEDECEGLACHQLCARGHGMGPGLTQCNCQSO 533
Db 472 LKRASTINMRRLFRSEDOSEIYDART-----ENQTCNNESEDOCMGPGRTMVCSCILH 524
Qy 534 FLRGQCEVECRYLOGLPREYVYNARHCLPCHPECOPONGSVTCFPEADQCACAHYKDP 593
Db 525 VPRGRCVASCNMLQGEPRADQVDRCYOCHQCELVQDTSLTLCYGGPRANCKCAHFQDG 584
Qy 594 PCVACRSPGVKPDLSYPIWKFPDEBAGACOPPCINCTHSCVDLDDKGC 642
Db 585 PCITPRCPHMLGDGDTL-TMKYADRMGOCQPOHONCTQCGSGPGLSGC 632

RESULT 9
09W6F6 PRELIMINARY: PRT: 1137 AA.
AC 09W6F6:
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Receptor tyrosine kinase (Fragment).
GN ERBB4.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
```

OX NCBI\_TaxId=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=HINDRAIN;  
RX MEDLINE=99263203; PubMed=10328884;  
RA Dixon M., Lumsden A.;  
RT "Distribution of neuregulin-1 (nrg1) and erbB4 transcripts in  
embryonic chick hindbrain."  
RL Mol. Cell. Neurosci. 13:237-258(1999).  
DR EMBL: AF121963; AAD31764.1; -  
DR HSSP: P11362; IFGK.  
DR InterPro: IPR000494; EGFR\_L\_domain.  
DR InterPro: IPR000719; Euk\_Pkinase.  
DR InterPro: IPR002174; Furin-like.  
DR InterPro: IPR001368; TNFR\_c6.  
DR InterPro: IPR001245; Tyr\_pkinase.  
DR InterPro: IPR004019; YLP\_motif.  
DR Pfam: PF00757; Furin-like; 1.  
DR Pfam: PF00069; pkinase; 1.  
DR Pfam: PF01030; Recep\_L\_domain; 1.  
DR Pfam: PF02757; YLP; 2.  
DR PRINTS: PR00109; TYRKINASE.  
DR PRODOM: PD000001; Euk\_pkinase; 1.  
DR SMART: SM00261; FU; 3.  
DR SMART: SM00219; TYRK; 1.  
DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE: PS00101; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE: PS00652; TNFR\_NGFR\_1; UNKNOWN\_1.  
KW kinase; tyrosine-protein kinase.  
FT NON\_TER  
SQ SEQUENCE 1137 AA; 127927 MW; 4D616436F87D84F CRC64;  
  
Query Match 24.7%; Score 1252.5; DB 13; Length 1137;  
Best Local Similarity 26.1%; Pred. No. 9.6e-89;  
Matches 296; Conservative 121; Mismatches 314; Indels 403; Gaps 21;  
  
QY 161 LCYDFTILMKDIFHNKNOALTLIDTNSRACHPCSPCKSGCESSEDSQSLTRTVG 220  
DB 3 LCFADTIHMODIVRNPMASNFTLVPTNGSSGCRHCKSCGTG-RCMGPTENHQTLTKIVC 61  
  
QY 221 AGGC-ARCKGFLPTDCCHEGCAAGCTGPKHSDCLACLFHNHSGICELHCPALVYINDTF 279  
DB 62 AEQCDGRCYGYSPVSDCCHEGCAAGCTGPKHSDCLACLFHNHSGICELHCPALVYINDTF 121  
  
QY 280 ESMNPBEGRYTFEGASCYACPYNYLSTDVGSCTLVCPHNEGVNAEDTOCECKSRCA 339  
DB 122 QLENNHNAKYTYGAFYKACPHNFY-VDSSCVRACPSSKMEV-ENGIKMKCKPCTDICP 179  
  
QY 340 RVCYGLGMEHLREYRAVYTSANIOEFAGCKIFGSLAFLPESFDGDPASNTAPLPEQLQV 399  
DB 180 KACGIGIGSLVSAQTVSSNIDKFINCTKINGNLIFLVTHGHPYHTIAINPEKINI 239  
  
QY 400 FETLEITGYLYISAMPDSDLPSYFONLQYIRGRIHNGAYSLTGLGISMGLRSLR 459  
DB 240 FQYREITGYLYNISWPMNTDFRYFSNLVYIGGRALYSGLSLILKQGGITSLQFOSLK 299  
  
QY 460 ELGSGALILHNHFLCEVHTVPMQDLFRNPQALHTANRPDEDECGGLACHOICARGH 519  
DB 300 QISAGNIYTDNSNLCTHYTNMTSLFSTPSQKTYIHNRKKAENCTAGMVCNEICSSDG 359  
  
QY 520 CMGPGPTCVNCSQFLRGQCEVRCVLAQLPREYVNAHCLPCHPECP-ONGSVTFG 578  
DB 360 CMGPGPTCVNCSQFLRGQCEVRCVLAQLPREYVNAHCLPCHPECP-ONGSVTFG 578  
  
QY 579 PEADQVACAHAHYKDPFVCANCPGSKVDPDLSYMPWKFPDEGACORPCINCHSCVDLD 638  
DB 420 PGPHCTKCFHFKDGPNVCEKCPDGLQGANF--IFKYADEDRCHPCHPNCTOGCGRPA 477  
  
QY 639 DKGC-----PAQR----- 647  
DB 478 SHDCIYPTWRTOSTLPQHAKRPLLAAGYIGLFIIVIMGLTFPAVYVRKSKIKKRALRRF 537

QY 648 -----ASPLTSQ----- 654  
DB 538 LETELVEPLTPSGAPNAQALRIKETELKRVKYLGSAGFGTVYKGAIWVPGETVKIPVA 597  
QY 655 -----NEDIGP----- 660  
DB 598 IKILETTGPKANVEFMDALIMASMDHPLVRLGLCPSPTIQLVTLMPHGLLDYVH 657  
QY 661 ----- 660  
DB 658 EHKDNISQDLLNMCVOYAKGMVLEBRRLVHRDLAANVLKSPNHVKTDFGLARILE 717  
QY 661 ----- 660  
DB 718 GDEKEYNADGKMDIKWMALECIHYRKFTHQSDVSYGVTIMEIMTEGKPYDGIPTREI 777  
QY 661 ----- 660  
DB 778 PDLEKGERLPDPICITIDVYVWVKWMIDADSRPKELAEFSRMARDPQRYLYIQG 837  
QY 661 -----ASPLDSTYRSLLEDMDMDLVDAEYLVPOGFCPPAPAGAGVHRRHS 714  
DB 838 DDRKRLPSNPDSKFPQNLDEDELEDMDADEYLV-QAFNIPPTIYSTRIRDSNRQF 896  
QY 715 STRSGGDLTLGLEPSEEEADRS--PLAP-SEGAGSDVFDGLGAAKGLQSLPTHDS 771  
DB 897 VYRQGYAAEGOV-PMYRABGCIPEAPVAGQATAEIFEDTCNGTLRKQVATLAKEDS 955  
QY 772 PLQRYSEPTVPLPS-----EIDGYAPLTCSPQRPYVQNPVQRPSPRGRPLPA 824  
DB 956 STQRYSDPTVFIPEYVIRGELDEGYVTPMRDKPKTYVLPVENPVSRRKNGDLOA- 1014  
QY 825 RPAGATLERPK-TLSPGKNGVYKDF-----AFGAVENPEYLTPOGGAAPQPPPP 875  
DB 1015 -----VDNPEYHNAQNPQPAEDEVYNEPLXIMTFATLENAEYL-----KNLPE 1060  
  
QY 876 AFSPAEDNLVYWDQDPPERGA--PPSTFKGTPY-----AENDEYL 913  
DB 1061 KAKKAFNDPQVWNHSLPRTSLQHPDYLQEXSTKYFKQNGRIRPIVANEPEYL 1114  
  
RESULT 10  
Q8R2X1 ID Q8R2X1 PRELIMINARY; PRT; 367 AA.  
AC Q8R2X1;  
DT 01-JUN-2002 (Tremblrel. 21, Created)  
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)  
DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
OS Hypothetical 40.2 kDa protein.  
OC Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxId=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: BC027080; AAH27080.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 367 AA; 40163 MW; 0BE03395F9E101B0 CRC64;  
  
Query Match 24.0%; Score 1217; DB 11; Length 367;  
Best Local Similarity 70.7%; Pred. No. 1.2e-86;  
Matches 241; Conservative 16; Mismatches 56; Indels 28; Gaps 5;  
  
QY 603 GVKPDLSTYMPWKPPD--EAGACPCPTNCT-----HSCVDLDDKGCAPQAPSLTS 653  
DB 31 GAKP-YDGIAPAREIPDLLEKGERLPQPICTIDVYMIWVKWMIDSECRP--RFRELVS 86  
  
QY 654 -----QNEDLGPAASPLDSTFYRSLLEDMDMDLVDAEYLVPOGFCPPD 698  
DB 87 EFSRMARDPQRYVVIQNEDLGPAASPLDSTFYRSLLEDMDMDLVDAEYLVPOGFCPPD 146

Oy	699	PAPGAGCVAHRRHHSSYRSGCGDGLTGLEPSEBEARSPPLAASEGASVPPGDGMGA	758
Dd	147	PALGTGSTARRHRSSSRKSGGGELTLGLEPSEEPRSPAASEGASVPFGDLAVGV	206
Oy	759	AKGLQSLPTHDPSPLOQRSEDPYPPLPSEFNDGYAAPLTCSPOEYVWQPDPVRQPSPRE	818
Dd	207	TKGILQSLSPHLSPLQLGRSDEPTLPLEPERFDGYAAPLACSPQEYVWQPDPVRQSLPRE	266
Oy	819	GPLLPAARPAGATLERPKYTLSPGKNKVYDVAFEGGAVENTPYLTPOGGAAPQHPHPAFS	878
Dd	267	GPPPIIRPAGATLERPKYTLSPGKNKVYDVAFEGGAVENTPYLARAGTAOPHPSPAFS	326
Oy	879	PAFDNLVYMDDPEERCAPSTFEKGPRTAAENPEYLIGLDIVPV	919
Dd	327	PAFDNLVYMDDONSSEQGPSPTEFGTPTAEENPEYLIGLDIVPV	367

ID	Q8WYV0	Q8WYV0	PRELIMINARY;	PRF;	412 AA.
AC	Q8WYV0				
DT	01-MAR-2002	(TREMBLrel. 20, Created)			
DT	01-MAR-2002	(TREMBLrel. 20, Last sequence update)			
DT	01-JUN-2002	(TREMBLrel. 21, Last annotation update)			
DE	Hypothetical 44.7 kDa protein.				
GN	p3659.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID:9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Huang Y., Zhou X.M., Zhang P.P., Jiang H.Q., Qin W.X., Zhao X.T.,				
RA	Wan D.F., Gu J.R.,				
RT	"Novel human cDNA clones with function of inhibiting cancer cell				
RT	growth."				
RL	Submitted (NCV-2000) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AF318349; AA155856.1; "				
DR	InterPro: IPR002048; EF-hand.				
DR	InterPro: IPR000719; Euk_kinase.				
DR	InterPro: IPR001245; Tyr_kinase.				
DR	InterPro: IPR004019; YLP_motif.				
DR	Pfam; PF00069; pkinase; 1.				
DR	Pfam; PF02757; TLP; 2.				
DR	ProDom; PD000001; Euk_pkinase; 1.				
DR	SMART; SM00219; TyrKc; 1.				
DR	PROSITE; PS00018; EF_HAND; UNKNOWN.1.				
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.				
KW	Hypothetical protein.				
SO	SEQUENCE 412 AA; 44702 MW; 034397FF5F27D2BC CAC64;				
	Query Match	23.1%	Score 1175.5;	DB 4;	Length 412;
	Best Local Similarity	64.6%;	Pred. No. 2.5e-83;		
	Matches 248;	Conservative 10;	Mismatches 45;	Indels 81;	Gaps 9;
QY	603 GVKPLDLSYMLMKRPD--EEGACGCPICNCT-----HSCVDLDDKCGPAQRASPLTS	653			
Db	31 GAKP-YDGIPIAREIPDLLEKGERLCPQPICTIDVYIMVKCMIDSECRP--RREELVS	86			
QY	654 -----QNEIDLGPASPLDSTFYRSLLEDDMDGLVDAEETIVPQGGFCPD	698			
Db	87 EFSRMARDPQRFFVYIQNEDLGPASPLDSTFYRSLLEDDMDGLVDAEETIVPQGGFCPD	146			
QY	699 PAPGAGGVHRRHSSSTRSGGDLITLGLFSESEEAAPSPAPLSPGACSDVFDGLGMA	758			
Db	147 PAPGAGGVHRRHSSSTRSGGDLITLGLFSESEEAAPSPAPLSPGACSDVFDGLGMA	206			
QY	759 AKGLQSLPTHPSPFLQRISEDPYVPLPSETDGYVAFLTCSPPQEVVNPDPVRPOPSPRE	818			
Db	207 AKGLQSLPTHPSPFLQRISEDPYVPLPSETDGYVAFLTCSPPQEVVNPDPVRPOPSPRE	266			
QY	819 GPLPAAPRAGATLERPKTLSPGKNGVAKVDFAFGCAVENPEYLTPQGGAAPD-----	871			

QY	DB	QY	DB
872	267	895	387
QPTSTTSTGCTRHOSGGLHPAPSGHLRQRTOSTWVNTCCCEPBGVARSBDVSSGREG	GLPLAARPAAGATLERPTLSPGKNQGVKQDVAFAGGAVENPEYLTPQGGALSPILLPSA	-----GAPSTFGKTPPAEN	LTSAGIKRWEGPPTTSRGTCAARN
895	387	909	410
-----GAPSTFGKTPPAEN	-----HPPPA---ESPFDNLVYWD---ODDPER	: :	: :

RESULT	ID	PRELIMINARY:	PRF:	527 AA.
090836	090836			
AC	090836;			
DT	01-NOV-1996 (TREMblrel. 01, Created)			
DT	01-NOV-1996 (TREMblrel. 01, Last sequence update)			
DT	01-JUN-2002 (TREMblrel. 21, last annotation update)			
DE	EGF/TCF-alpha receptor precursor.			
CN	C-ERBB.			
OS	Gallus gallus (Chicken).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;			
OC	Gallus.			
OX	NCBI_TaxID=9031;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92123214; PubMed=1732751;			
RA	Flickinger T.W., Mahle N.J., Kung H.-J.;			
RT	"An alternatively processed mRNA from the avian c-erbB gene encodes a			
RT	soluble, truncated form of the receptor that can block ligand-			
RT	dependent transformation."			
RL	Mol. Cell. Biol. 12:883-893(1992).			
DR	EMBL; M77637; AAA48759.1; -			
DR	InterPro; IPR000494; EGFR_L_domain.			
DR	InterPro; IPR002174; Furin-like.			
DR	Pfam; PF00757; Furin-like; 1.			
DR	Pfam; PF01030; Recep_L_domain; 2.			
DR	SMART; SMO0261; FU; 2.			
KM	Receptor; Signal.			
FT	SIGNAL	1	28	POTENTIAL.
FT	CHAIN	29	527	EGF/TCF-ALPHA RECEPTOR.
SO	SEQUENCE	527 AA;	58353 MW;	764564ABCC095298 CRC64;
Query Match		22.7%;	Score 1155;	DB 13; Length 527;
Best Local Similarity		44.0%;	Pred. No. 1.4e-81;	
Matches 227;	Conservative	87;	Mismatches 182;	Indels 20; Gaps
QY	11 LLLALLPFGAASST-----QVCTGTDMLKRLRSPETHIDMLRNHLKQSGOVGNILETYL	65		
Db	20 LLLLLLGRVALCSANEEKKVCGGTNNKLTQLGSHVDHFTSLQRMNNEEVALSNLEITY	79		
QY	66 PTNNSLSFLDIOEYQGYVLLAHNOVROYPLQRLRYVGTQLFEDNVALAVLDNGDPLNN	125		
Db	80 EHNRLFLFKLTIOEAGVGLIALNVDVPLENGLIINGVALYDNSFALVLSNYH-MNK	138		
QY	126 TTPVTGASPPGRLRLELTLKGLVLIQNPOLCYODTILKMDIFHNKNNQLATLTLD	185		
Db	139 TQ-----GLRELPMKRLSELILNGVKISINPRLCMMDVLLMNDIITSRK-PLTVLD	189		
QY	186 -TNRSRACHPSCPMCKGSRGSESEDDQSLRTVCAGGCA-RCKGRLPTDCHEBOCAAG	243		
Db	190 FASNLSQCKPKHNPCTEDHCGAGSNGQTLTYVCAQCCSGRCGRKVPSPDCCHHQCAAG	249		
QY	244 CTGRHSDCLACLPHNHSIGIELHCPLALVYVNTDTFESMPNDEGRYTFGASCVTPAPYNY	303		
Db	250 CTGREGSCICLACRFKRDATCKDTCPLVLYNPTTYQMDVNPDEGKYSFGATCVRECPHNY	309		
QY	304 LSTVGSTLVCPLRLNQEVTAEQGNOREKSKSCARVCGLSGEMHLEFVRAVANSQAE	363		
Db	310 VYIDHGSVCVNSCNMDTYEV--EENGRRCKKCKDGLGKSCVNGSIGELGLISINNTNIDS	368		
QY	364 FAGCKKIFGSLAFIPESHDDGPASNTAPLQEOLOVEFTLEIIGLYLISAMPDLPDLS	423		



[illegible]

Search completed: January 13, 2003, 14:48:25  
Job time : 51.2593 secs

**THIS PAGE BLANK (USPTO)**







```

Db 361 IOEFAGCKKIFGSLAFPEPSFGDDPASNTAPLQBPOLQVFEETLEITGYLTISAMPDSL 420
Qy 421 DLSVFQNLQVIRGRIHNGAYSILTLQGLISWLGRLSLELSSGLALIHNNHLCFVHTV 480
Db 421 DLSVFQNLQVIRGRIHNGAYSILTLQGLISWLGRLSLELSSGLALIHNNHLCFVHTV 480
Qy 481 PMDOLFNRPHQALLHTANPEDECVGEGLAQHOLCARHCWGPGPTQCNCQSFILRGDGC 540
Db 481 PMDOLFNRPHQALLHTANPEDECVGEGLAQHOLCARHCWGPGPTQCNCQSFILRGDGC 540
Qy 541 VEECRVLOGLPREYVNAHRLCPHRECOPONSVCFCGEADQCVACAHYKPPFCVAC 600
Db 541 VEECRVLOGLPREYVNAHRLCPHRECOPONSVCFCGEADQCVACAHYKPPFCVAC 600
Qy 601 PSGVPRDLSYMPIMKFPDEGACOPCPINCSTHSCVDLDKGCAPABQASPLSIISAVVG 660
Db 601 PSGVPRDLSYMPIMKFPDEGACOPCPINCSTHSCVDLDKGCAPABQASPLSIISAVVG 660
Qy 654 ----- 653
Db 661 ILLVVLGVVFGILIKRQOKIRKYMRLLOETELVEPLPSGAMPNOQRILKETEL 720
Qy 654 ----- 653
Db 721 RKVYVLSGAFGTYYKGIWIPGENVKIPVALKRENTSPRANKELIDEATVMAGVSP 780
Qy 654 ----- 653
Db 781 YVSRLLIGLITSTVOLYQMLPYGCLLDHVENRGLSQDLLNMCQIAKMSYLEYR 840
Qy 654 ----- 653
Db 841 LVHRLAARNVLYKSPNHVKITDFGLARLLIDETEVHADGSKVPIKMALESILRRFT 900
Qy 654 ----- 653
Db 901 HOSDWSYGVTVWELMTGSAKPYDGIAPAREIPDLLEKGERLQPPICITIDVYIMVKKM 960
Qy 654 ----- QNEDLGASPLDSTFYRSLLEDDMDGLVDA 684
Db 961 IDSECRPRELSEFSRMAPRPFVVYQNLNEDLGASPLDSTFYRSLLEDDMDGLVDA 1020
Qy 685 EEYLVPOOGFFCDDPARAGAGVHHHRH 712
Db 1021 EEYLVPOOGFFCDDPARAGAGVHHHRH 1048

```

# RESULT 3

US-08-467-083-68

Sequence 68, Application US/08467083

Patent No. 5726023

GENERAL INFORMATION:

APPLICANT: Cheever, Martin A.

APPLICANT: Dsils, Mary L.

TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/NEU PROTEIN

TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE

TITLE OF INVENTION: HER-2/NEU ONCOGENE IS ASSOCIATED

NUMBER OF SEQUENCES: 68

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed and Berry

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: US

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/467, 083

FILING DATE: 06-JUN-1995

```

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/414,417
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
TELEX: 3723836 SEEDANBERRY
INFORMATION FOR SEQ. ID NO.: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-467-083-68

```

Query Match 95.3%; Score 3767; DB 1; Length 1255;  
Best Local Similarity 67.8%; Pred. No. 7.8e-317;  
Matches 711; Conservative 0; Mismatches 1; Indels 336; Gaps 1;

```

Qy 1 MELAALCRWGLLALIPGAASVOTGDMKRLRPASPEHLDMLRHLYOGCQVVGML 60
Db 1 MELAALCRWGLLALIPGAASVOTGDMKRLRPASPEHLDMLRHLYOGCQVVGML 60
Qy 61 ELTYLPTNASLFLDIOGVGVLIHNOVQVPLQRLRIYRGTLQFEDNALAVLDNG 120
Db 61 ELTYLPTNASLFLDIOGVGVLIHNOVQVPLQRLRIYRGTLQFEDNALAVLDNG 120
Qy 121 DPLNTPYVYGASPGGLRELQRLSLTEILKGVLIQRPOLCYODTILMKDIFHKNNOLA 180
Db 121 DPLNTPYVYGASPGGLRELQRLSLTEILKGVLIQRPOLCYODTILMKDIFHKNNOLA 180
Qy 181 LFLIDTNRBRACHPCSPMKGSRGWESSEDOSLTRVACAGGACARCKPLTDCHEQC 240
Db 181 LFLIDTNRBRACHPCSPMKGSRGWESSEDOSLTRVACAGGACARCKPLTDCHEQC 240
Qy 241 AAGCTGPKHSDCLACLFHNSGICELCPALVTYNTDFESMPNREGRYTFGASCYTAC 300
Db 241 AAGCTGPKHSDCLACLFHNSGICELCPALVTYNTDFESMPNREGRYTFGASCYTAC 300
Qy 301 YNLTSDVSGCTLVCPPLHNOEYTAEDGTORCEKSPCARVYCYGLAMEHLEVRVATSN 360
Db 301 YNLTSDVSGCTLVCPPLHNOEYTAEDGTORCEKSPCARVYCYGLAMEHLEVRVATSN 360
Qy 361 IOEFAGCKKIFGSLAFPEPSFGDDPASNTAPLQBPOLQVFEETLEITGYLTISAMPDSL 420
Db 361 IOEFAGCKKIFGSLAFPEPSFGDDPASNTAPLQBPOLQVFEETLEITGYLTISAMPDSL 420
Qy 421 DLSVFQNLQVIRGRIHNGAYSILTLQGLISWLGRLSLELSSGLALIHNNHLCFVHTV 480
Db 421 DLSVFQNLQVIRGRIHNGAYSILTLQGLISWLGRLSLELSSGLALIHNNHLCFVHTV 480
Qy 481 PMDOLFNRPHQALLHTANPEDECVGEGLAQHOLCARHCWGPGPTQCNCQSFILRGDGC 540
Db 481 PMDOLFNRPHQALLHTANPEDECVGEGLAQHOLCARHCWGPGPTQCNCQSFILRGDGC 540
Qy 541 VEECRVLOGLPREYVNAHRLCPHRECOPONSVCFCGEADQCVACAHYKPPFCVAC 600
Db 541 VEECRVLOGLPREYVNAHRLCPHRECOPONSVCFCGEADQCVACAHYKPPFCVAC 600
Qy 601 PSGVPRDLSYMPIMKFPDEGACOPCPINCSTHSCVDLDKGCAPABQASPLSIISAVVG 660
Db 601 PSGVPRDLSYMPIMKFPDEGACOPCPINCSTHSCVDLDKGCAPABQASPLSIISAVVG 660
Qy 654 ----- 653
Db 661 ILLVVLGVVFGILIKRQOKIRKYMRLLOETELVEPLPSGAMPNOQRILKETEL 720
Qy 654 ----- 653

```

Db 721 RKVVLGSGAGFYKGIWIPDGENVKIPVAKVRENTSPKANKELIDEAYVWAGVSP 780  
QY 654 ----- 653  
Db 781 YVSRLLGICLSTVQVLTQMLPFGCLLDHVENRGRUGSODLLWMCQIAKMSYLEDV 840  
QY 654 ----- 653  
Db 841 LVHRLAARNLVKSPNHVKITDFGLARLDIDETEXHADGKVPKMALESILRRFT 900  
QY 654 ----- 653  
Db 901 HQSDWWSYGVTVWELMTFGAKPYDGIAPARELPDLEKGERLPDPICITIDVYIMVWCM 960  
QY 654 ----- 653  
Db 961 IDSECRPRFRELVESEFSMARDPQRFVYIQNEDLGPSPLDSTFYRSLLEDMDGLDVA 1020  
QY 685 EEYLVPOGFFCPDPAPAGAGMVHRRH 712  
Db 1021 EEYLVPOGFFCPDPAPAGAGMVHRRH 1048

RESULT 4  
US-08-414-417B-68  
; Sequence 68, Application US/08414417B  
; Patent No. 5801005  
; GENERAL INFORMATION:  
; APPLICANT: Cheever, Martin A.  
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN  
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
; NUMBER OF SEQUENCES: HER-2/neu ONCOGENE IS ASSOCIATED  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed and Berry LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: US  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/414,417B  
; FILING DATE: 31-MAR-1995  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharkey, Richard G.  
; REGISTRATION NUMBER: 32,629  
; REFERENCE/DOCKET NUMBER: 920010, 448C2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 68:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1255 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; US-08-414-417B-68

Query Match 95.3%; Score 3767; DB 1; Length 1255;  
Best Local Similarity 67.8%; Pred. No. 7.8e-317;  
Matches 711; Conservative 0; Mismatches 1; Indels 336; Gaps 1;

QY 1 MELALRCRWGILLALLPGASTOVCTGDKLRLPASPEHIMLRILYOGCVOYOGNT 60  
Db 1 MELALRCRWGILLALLPGASTOVCTGDKLRLPASPEHIMLRILYOGCVOYOGNT 60  
QY 61 ELTYLPTNASLSFLQDIOEVGVVLIANOVROYPLQRLIRVGTOLFEEDYALAVLDNG 120

Db 61 ELTYLPTNASLSFLQDIOEVGVVLIANOVROYPLQRLIRVGTOLFEEDYALAVLDNG 120  
QY 121 DPLNNTPTVTGASPGGLRELOLRSLTEILKGVLIQORNPOLCYODTIIKMDIFKNNOLA 180  
Db 121 DPLNNTPTVTGASPGGLRELOLRSLTEILKGVLIQORNPOLCYODTIIKMDIFKNNOLA 180  
QY 181 LTLIDITNRSRACHPCSPKCKSRMGSESSDQSLTFTVAGGACRCKGPLPTDCHBQC 240  
Db 181 LTLIDITNRSRACHPCSPKCKSRMGSESSDQSLTFTVAGGACRCKGPLPTDCHBQC 240  
QY 241 AAGCTGPKHSCLACLFHNSGICELHCPALVYNTDTFESMPNPEGRYTFGASCVACP 300  
Db 241 AAGCTGPKHSCLACLFHNSGICELHCPALVYNTDTFESMPNPEGRYTFGASCVACP 300  
QY 301 YNYLSTDVGSCTLVCPPLHNOEVTAEADGTQRCCKSKPCARVCYIGMEHLREVAVTISAN 360  
Db 301 YNYLSTDVGSCTLVCPPLHNOEVTAEADGTQRCCKSKPCARVCYIGMEHLREVAVTISAN 360  
QY 361 IQEFAGCKKIRGSLAFLEPSEFDGDPASNTAPLOPEQLOVEFTELEITGYLISAMPDLP 420  
Db 361 IQEFAGCKKIRGSLAFLEPSEFDGDPASNTAPLOPEQLOVEFTELEITGYLISAMPDLP 420  
QY 421 DLSVFQNLQVIRGRILHNGAVSLTLQGLISWLGRLSRLRELSGLALIHNTHCIFYHTV 480  
Db 421 DLSVFQNLQVIRGRILHNGAVSLTLQGLISWLGRLSRLRELSGLALIHNTHCIFYHTV 480  
QY 481 PMDOLFNRPHQALLHTANRPDECEVGEGLACHOLCARGHCMPPTQCVNCSQFLRQEC 540  
Db 481 PMDOLFNRPHQALLHTANRPDECEVGEGLACHOLCARGHCMPPTQCVNCSQFLRQEC 540  
QY 541 VEECRVLQGLREYVNAHCLPCHECPONGSVTCGPEADQCVACAHYDPPFCVARC 600  
Db 541 VEECRVLQGLREYVNAHCLPCHECPONGSVTCGPEADQCVACAHYDPPFCVARC 600  
QY 601 PSQVVKPDLSTYPIKFPDEBEGACOPCPTNCTHSCVDDDKCPAEQASPLTS----- 653  
Db 601 PSQVVKPDLSTYPIKFPDEBEGACOPCPTNCTHSCVDDDKCPAEQASPLTS----- 653  
QY 654 ----- 653  
Db 661 ILVVVLGVVFGILLKRRQOKIRRYTMRRLQETELVEPLTPSGAMPNOAMRLKTEL 720  
QY 654 ----- 653  
Db 721 RKVVLGSGAGFYKGIWIPDGENVKIPVAKVRENTSPKANKELIDEAYVWAGVSP 780  
QY 654 ----- 653  
Db 781 YVSRLLGICLSTVQVLTQMLPFGCLLDHVENRGRUGSODLLWMCQIAKMSYLEDV 840  
QY 654 ----- 653  
Db 841 LVHRLAARNLVKSPNHVKITDFGLARLDIDETEXHADGKVPKMALESILRRFT 900  
QY 654 ----- 653  
Db 901 HQSDWWSYGVTVWELMTFGAKPYDGIAPARELPDLEKGERLPDPICITIDVYIMVWCM 960  
QY 654 ----- 653  
Db 961 IDSECRPRFRELVESEFSMARDPQRFVYIQNEDLGPSPLDSTFYRSLLEDMDGLDVA 1020  
QY 685 EEYLVPOGFFCPDPAPAGAGMVHRRH 712  
Db 1021 EEYLVPOGFFCPDPAPAGAGMVHRRH 1048

RESULT 5  
US-08-486-348A-68  
; Sequence 68, Application US/08486348A  
; Patent No. 5846538  
; GENERAL INFORMATION:

```

APPLICANT: Cheever, Martin A.
APPLICANT: Disis, Mary L.
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
NUMBER OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,348A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C6
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-486-348A-68

Query Match          95.3%  Score 3767; DB 2; Length 1255;
Best Local Similarity 67.8%  Pred. No. 7.8e-317;
Matches 711; Conservative 0; Mismatches 1; Indels 336; Gaps 1;

```

```

|||||
Db 481 PMDOLFNRPHQALHTANRPNDECVGEGIALCHQLCARGCHWGPRPTQVCNCSQFLRGQRC 540
|||||
Qy 541 VEECRVYLOGLPREYVNAHRLCPHCECOPQNSVTCFGEADQCVAAHYKDPPECVAC 600
|||||
Db 541 VEECRVYLOGLPREYVNAHRLCPHCECOPQNSVTCFGEADQCVAAHYKDPPECVAC 600
Qy 601 PSQVKRDLSTYMPKPFDEGACQPCPINCISCVLDLCKGCPAEGRASPLTS----- 653
|||||
Db 601 PSQVKRDLSTYMPKPFDEGACQPCPINCISCVLDLCKGCPAEGRASPLTSISAVVG 660
Qy 654 ----- 653
Db 661 ILLVYVGLVYFGLIKRRQOKIRKTMRLLOTELVEPLTDSGAMPNOQMRILKETEL 720
Qy 654 ----- 653
Db 721 RKVKVLGSGAFGVYKGIWIPGENVKIPVAIKVLENTSPKANKELDEAVVMAGVQSP 780
Qy 654 ----- 653
Db 781 YVSRLLGICLTSTVOLVTLMPYGLLDHVRNRRGLSQDILLNMQOIAKGMSTLEBYR 840
Qy 654 ----- 653
Db 841 LVHRDLAANLVKSPNHVKITDFGLARLLDIDETEFHADGGKVPFKMALESILRRRT 900
Qy 654 ----- 653
Db 901 HQSDVMSGVYVWELMTFGAKPYDCIPAREIPDLLEKGERLPORPCTIDVYIMVYKWM 960
Qy 654 ----- 1021
Db 961 IDSECRPRELIVSEFSRNARDPQRFVTLNEDLGPASPLDSYFRSLLEDMDGLVDA 1020
Qy 685 EEYLVPOQGFPCPDPAAGAGVHHRR 712
Db 1021 EEYLVPOQGFPCPDPAAGAGVHHRR 1048

RESULT 6
US-08-486-545B-68
Sequence 68, Application US/08468545B
Patent No. 5876712
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,545B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C5
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900

```

TELEFAX: (206) 682-6031  
 INFORMATION FOR SEQ ID NO: 68:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1255 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 US-08-468-545B-68

Query Match 95.3%; Score 3767; DB 2; Length 1255;  
 Best Local Similarity 67.8%; Pred. No. 7, 8e-317;  
 Matches 711; Conservative 0; Mismatches 1; Indels 336; Gaps 1;

```

OY 1 MELAALCRWGLLLALLPPGAASTOVCTGTDKRLPASPETHLMDLRLHYGCGVQVGNL 60
DB 1 MELAALCRWGLLLALLPPGAASTOVCTGTDKRLPASPETHLMDLRLHYGCGVQVGNL 60
OY 61 ELTYLPTNASLSFLQDIOEVGYVLAHNOVROVPLQRLRIVRGTOLEFEDNYALAVDNG 120
DB 61 ELTYLPTNASLSFLQDIOEVGYVLAHNOVROVPLQRLRIVRGTOLEFEDNYALAVDNG 120
OY 121 DPLNNTPTVPGASPGGLREQLRSLEILKGVLIQNPOLCYQDITLMKDIFKKNOLA 180
DB 121 DPLNNTPTVPGASPGGLREQLRSLEILKGVLIQNPOLCYQDITLMKDIFKKNOLA 180
OY 121 DPLNNTPTVPGASPGGLREQLRSLEILKGVLIQNPOLCYQDITLMKDIFKKNOLA 180
DB 121 DPLNNTPTVPGASPGGLREQLRSLEILKGVLIQNPOLCYQDITLMKDIFKKNOLA 180
OY 181 LTLIDTNRSRACHPCSPCKSGRCWGESSEDCQSLRTVCAGGACRCKGPLPTDCHEQC 240
DB 181 LTLIDTNRSRACHPCSPCKSGRCWGESSEDCQSLRTVCAGGACRCKGPLPTDCHEQC 240
OY 241 AAGCTGKHSDDLACLFHNSGICELCPALVYNTDTFESMPREERYTFGASCYACR 300
DB 241 AAGCTGKHSDDLACLFHNSGICELCPALVYNTDTFESMPREERYTFGASCYACR 300
OY 301 YNLTSTVSGCTVLCPLHNOEYTAEDGTORCEKSKPCARCYGLGMEHLREVAAYTSAN 360
DB 301 YNLTSTVSGCTVLCPLHNOEYTAEDGTORCEKSKPCARCYGLGMEHLREVAAYTSAN 360
OY 361 IOEFAGCKKIFGSLAFPLESFSDGDPASNTAPLOEQLQVEETLEITGYLYISAMPDLP 420
DB 361 IOEFAGCKKIFGSLAFPLESFSDGDPASNTAPLOEQLQVEETLEITGYLYISAMPDLP 420
OY 421 DLSYFOQLQVIRGILNGLAYSLTLOGLSWGLSLRELGSGLAIHNHTHLCFHTV 480
DB 421 DLSYFOQLQVIRGILNGLAYSLTLOGLSWGLSLRELGSGLAIHNHTHLCFHTV 480
OY 481 PMDLFNNPQALHTANREDECEVSGELACHQLCARGHCGPPTQCVNCSQFLRGQEC 540
DB 481 PMDLFNNPQALHTANREDECEVSGELACHQLCARGHCGPPTQCVNCSQFLRGQEC 540
OY 541 VEECRVLQGLPREYVNAHCLPCHECQPOGNGSYTCGPEADQCVACAHTKDPFCVARC 600
DB 541 VEECRVLQGLPREYVNAHCLPCHECQPOGNGSYTCGPEADQCVACAHTKDPFCVARC 600
OY 601 PSGVKPRLSTMPKPFDEBACOPCPINCHSCVDLDDKCPAEOASPLTS----- 653
DB 601 PSGVKPRLSTMPKPFDEBACOPCPINCHSCVDLDDKCPAEOASPLTS----- 653
OY 654 ----- 653
DB 661 ILVVVVLGVFGLIKRRQCKIRKTYMRLLQETELVEPLTPSGAMPNQAMRLKETEL 720
OY 654 ----- 653
DB 721 RKVVLGSAFGVYKGIWIPDGENKIPVAIKVIRENTSPKANKELIDEAYVAVAGVSP 780
OY 654 ----- 653
DB 781 YVSRLLGICLTSTVQVLTQMPYGLLDHVRKNGRLGSDLLMWCQIAKMSYLEDVR 840
OY 654 ----- 653
DB 841 LVHEDLAARNLVKSPNHVKITDGLARLLDIDETFYHAOGKXVPIKMALESTLRRFT 900
OY 654 ----- 653

```

```

DB 901 HQSDVMSYGVTVNMLTFGAKPYDGIAPREIPDLLEKGERLPQPPICITIDVYIMKCMW 960
OY 654 -----ONEDLGASPLDSTFYRSLLEDDMGDLVDA 684
DB 961 IDSECRPRRELYSEFSRMAKDPQRFVYIONEDLGASPLDSTFYRSLLEDDMGDLVDA 1020
OY 685 EEXLVPOQGFCCPDPAAGAGVHHRR 712
DB 1021 EEXLVPOQGFCCPDPAAGAGVHHRR 1048

```

RESULT 7  
 US-08-466-680B-68  
 Sequence 68, Application US/08466680B  
 Patent No. 6075122

GENERAL INFORMATION:

APPLICANT: Cheever, Martin A.

TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN

TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
 NUMBER OF SEQUENCES: 69

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed and Berry LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: US

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/466, 680B

FILING DATE: 06-JUN-1995

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Sharkey, Richard G.

REGISTRATION NUMBER: 32,629

REFERENCE/DOCKET NUMBER: 920010, 448C4

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 68:

SEQUENCE CHARACTERISTICS:

LENGTH: 1255 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-08-466-680B-68

Query Match 95.3%; Score 3767; DB 3; Length 1255;  
 Best Local Similarity 67.8%; Pred. No. 7, 8e-317;  
 Matches 711; Conservative 0; Mismatches 1; Indels 336; Gaps 1;

```

OY 1 MELAALCRWGLLLALLPPGAASTOVCTGTDKRLPASPETHLMDLRLHYGCGVQVGNL 60
DB 1 MELAALCRWGLLLALLPPGAASTOVCTGTDKRLPASPETHLMDLRLHYGCGVQVGNL 60
OY 61 ELTYLPTNASLSFLQDIOEVGYVLAHNOVROVPLQRLRIVRGTOLEFEDNYALAVDNG 120
DB 61 ELTYLPTNASLSFLQDIOEVGYVLAHNOVROVPLQRLRIVRGTOLEFEDNYALAVDNG 120
OY 121 DPLNNTPTVPGASPGGLREQLRSLEILKGVLIQNPOLCYQDITLMKDIFKKNOLA 180
DB 121 DPLNNTPTVPGASPGGLREQLRSLEILKGVLIQNPOLCYQDITLMKDIFKKNOLA 180
OY 121 DPLNNTPTVPGASPGGLREQLRSLEILKGVLIQNPOLCYQDITLMKDIFKKNOLA 180
DB 121 DPLNNTPTVPGASPGGLREQLRSLEILKGVLIQNPOLCYQDITLMKDIFKKNOLA 180
OY 181 LTLIDTNRSRACHPCSPCKSGRCWGESSEDCQSLRTVCAGGACRCKGPLPTDCHEQC 240
DB 181 LTLIDTNRSRACHPCSPCKSGRCWGESSEDCQSLRTVCAGGACRCKGPLPTDCHEQC 240

```

QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESMPNEGRTTFCASCVTACP 300  
| | | | |  
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESMPNEGRTTFCASCVTACP 300  
QY 301 YNYLSTDVSGCTLYCPLHNOEYTAEDGTORCEKSKPCARVCYGLGMEHLREAVTSAN 360  
| | | | |  
DB 301 YNYLSTDVSGCTLYCPLHNOEYTAEDGTORCEKSKPCARVCYGLGMEHLREAVTSAN 360  
QY 361 IOEFAGCKKIFGSLAFLEPSEFDGDPASNTAPLQBPOLQVFFLEETITGLYLSAMPDSL 420  
| | | | |  
DB 361 IOEFAGCKKIFGSLAFLEPSEFDGDPASNTAPLQBPOLQVFFLEETITGLYLSAMPDSL 420  
QY 421 DLSTVQNLQVIRGRIHNGAYSLTQGLGISWLGRLSRLREGSGALIHNTHLCEVHTV 480  
| | | | |  
DB 421 DLSTVQNLQVIRGRIHNGAYSLTQGLGISWLGRLSRLREGSGALIHNTHLCEVHTV 480  
QY 481 PWDQLFRNHQALHTANRPEDECVGEGLACHQCLARGCHWGPGPTQCVNSQPLRGQEC 540  
| | | | |  
DB 481 PWDQLFRNHQALHTANRPEDECVGEGLACHQCLARGCHWGPGPTQCVNSQPLRGQEC 540  
QY 541 VEECVNLOGLPREYVNAHCLPCHECOPONGSVTCFGEADQCVACAHYKDPFCVARC 600  
| | | | |  
DB 541 VEECVNLOGLPREYVNAHCLPCHECOPONGSVTCFGEADQCVACAHYKDPFCVARC 600  
QY 601 PSGVPRDLSTYMPIMKFPDEBEGACQPCPICTHSCVDLDDKGCAPORASPLTS----- 653  
| | | | |  
DB 601 PSGVPRDLSTYMPIMKFPDEBEGACQPCPICTHSCVDLDDKGCAPORASPLTS----- 653  
QY 654 ----- 653  
DB 654 ----- 653  
QY 654 ILLVVVLGVFEGILIKRROQKIRKTYMRLLQETELVEPLPFGAMPNQAKRLKETEL 720  
| | | | |  
DB 654 ILLVVVLGVFEGILIKRROQKIRKTYMRLLQETELVEPLPFGAMPNQAKRLKETEL 720  
QY 721 RKVKVLSGAGCTYKGIWIPDGENVKIPVAIKVLRNTPSPRANKETIDEAYVAVGVP 780  
| | | | |  
DB 721 RKVKVLSGAGCTYKGIWIPDGENVKIPVAIKVLRNTPSPRANKETIDEAYVAVGVP 780  
QY 781 YVSRLLGICLSTVOLVQLMPYGLLDHVRNKRGLSODLLNMCQIAKMSYLEDYR 840  
| | | | |  
DB 781 YVSRLLGICLSTVOLVQLMPYGLLDHVRNKRGLSODLLNMCQIAKMSYLEDYR 840  
QY 841 LVHRDLAARNVLVKSNNHVKITDFGLARLLDIDETEHADGSKVPIKMALESILRRRT 900  
| | | | |  
DB 841 LVHRDLAARNVLVKSNNHVKITDFGLARLLDIDETEHADGSKVPIKMALESILRRRT 900  
QY 901 HOSDVSYSVTVWELMEFGAPRYDIPAREIPDLLEKGERLPQPICTIDYVMIMVKCM 960  
| | | | |  
DB 901 HOSDVSYSVTVWELMEFGAPRYDIPAREIPDLLEKGERLPQPICTIDYVMIMVKCM 960  
QY 961 IDSECRPRFRELVSFESKARDPQRFVYIIONEDLGPASPLDSTFYSLLEDDMDGLVDA 1020  
| | | | |  
DB 961 IDSECRPRFRELVSFESKARDPQRFVYIIONEDLGPASPLDSTFYSLLEDDMDGLVDA 1020  
QY 1021 EBYLVPOQGFECRDPAPGAGGVNHRHR 1048  
| | | | |  
DB 1021 EBYLVPOQGFECRDPAPGAGGVNHRHR 1048

RESULT 8  
US-08-484-438-8

: Sequence 8, Application US/08484438  
: Patent No. 5811098  
: Patent No. 5811098 5780031  
: GENERAL INFORMATION:  
: APPLICANT: PLOWMAN, Gregory D.  
: APPLICANT: Culouscou, Jean-Michel  
: APPLICANT: Shoyab, Mohammed  
: APPLICANT: Siegall, Clay B.  
: APPLICANT: Hellstr m, Ingegerd  
: APPLICANT: Hellstr m, Karl E.  
: TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE  
: NUMBER OF SEQUENCES: 42  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Pennle & Edmonds

STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,438  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 08/323,442  
FILING DATE: 14-OCT-1994  
APPLICATION NUMBER: US 08/150,704  
FILING DATE: 10-NOV-1993  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/981,165  
FILING DATE: 24-NOV-1992  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-230  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1255 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-484-438-8

Query Match 94.3%; Score 3727; DB 2; Length 1255;  
Best Local Similarity 67.5%; Pred. No. 2.2e-313;  
Matches 707; Conservative 0; Mismatches 5; Indels 336; Gaps 1;  
QY 1 MELALCRWGLDALLPPGAASVQCTGTDKMLRLPASPEHIDMLRLHYGCGVYQGNL 60  
| | | | |  
DB 1 MELALCRWGLDALLPPGAASVQCTGTDKMLRLPASPEHIDMLRLHYGCGVYQGNL 60  
QY 61 ELTYIPTNASLSFLDDIDFVQGVYLIANNQVROVPLQRLRIYRGTLQFEDNYALAVLNG 120  
| | | | |  
DB 61 ELTYIPTNASLSFLDDIDFVQGVYLIANNQVROVPLQRLRIYRGTLQFEDNYALAVLNG 120  
QY 121 DPLNNTTPTVGTASPGGLRELQRLSTETIKGVLTORNDPOLCYODTILMKDIFRNNOLA 180  
| | | | |  
DB 121 DPLNNTTPTVGTASPGGLRELQRLSTETIKGVLTORNDPOLCYODTILMKDIFRNNOLA 180  
QY 181 LTLIDTNSRACHPCSPCKSRGCESSDQSLRTVCAAGCARGKPLPTDCCHQC 240  
| | | | |  
DB 181 LTLIDTNSRACHPCSPCKSRGCESSDQSLRTVCAAGCARGKPLPTDCCHQC 240  
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESMPNEGRTTFCASCVTACP 300  
| | | | |  
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESMPNEGRTTFCASCVTACP 300  
QY 301 YNYLSTDVSGCTLYCPLHNOEYTAEDGTORCEKSKPCARVCYGLGMEHLREAVTSAN 360  
| | | | |  
DB 301 YNYLSTDVSGCTLYCPLHNOEYTAEDGTORCEKSKPCARVCYGLGMEHLREAVTSAN 360  
QY 361 IOEFAGCKKIFGSLAFLEPSEFDGDPASNTAPLQBPOLQVFFLEETITGLYLSAMPDSL 420  
| | | | |  
DB 361 IOEFAGCKKIFGSLAFLEPSEFDGDPASNTAPLQBPOLQVFFLEETITGLYLSAMPDSL 420

```

QY 421 DLSVFQNLQVIRGRILHNGAYSLTLOGLSIWLGLSLRLSGSLALIHNTHLCEVHTV 480
DLSVFQNLQVIRGRILHNGAYSLTLOGLSIWLGLSLRLSGSLALIHNTHLCEVHTV 480
DLSVFQNLQVIRGRILHNGAYSLTLOGLSIWLGLSLRLSGSLALIHNTHLCEVHTV 480
QY 481 PMDQLEFRNHQALHTANRDECEVGEGLACHOLCARHGHWGPPQCVNCSQFLRGQC 540
PMDQLEFRNHQALHTANRDECEVGEGLACHOLCARHGHWGPPQCVNCSQFLRGQC 540
PMDQLEFRNHQALHTANRDECEVGEGLACHOLCARHGHWGPPQCVNCSQFLRGQC 540
QY 541 VEECRVLOGLPREYVNAHRLCPHPECOPQNSVTCGPEADOCVACAHKDDPPFCVAC 600
VEECRVLOGLPREYVNAHRLCPHPECOPQNSVTCGPEADOCVACAHKDDPPFCVAC 600
VEECRVLOGLPREYVNAHRLCPHPECOPQNSVTCGPEADOCVACAHKDDPPFCVAC 600
QY 601 PSGVKPDLSTYMPIMKFPDEGACQPCPINCTHSCVDLDKGCRAEORASPLTISVAVG 660
PSGVKPDLSYMPIMKFPDEGACQPCPINCTHSCVDLDKGCRAEORASPLTISVAVG 660
PSGVKPDLSYMPIMKFPDEGACQPCPINCTHSCVDLDKGCRAEORASPLTISVAVG 660
QY 654 ----- 653
661 ILVVLVGVVFGILIRROOKIRKYMRLLOETELVEPLTPSGAMPNOQMRLKETEL 720
661 ILVVLVGVVFGILIRROOKIRKYMRLLOETELVEPLTPSGAMPNOQMRLKETEL 720
661 ILVVLVGVVFGILIRROOKIRKYMRLLOETELVEPLTPSGAMPNOQMRLKETEL 720
QY 654 ----- 653
721 RKVKVLGSAFGTVYKGIWIPDGENKIPVAIVLRENTSPRANKELDEAYVAVGSP 780
RKVKVLGSAFGTVYKGIWIPDGENKIPVAIVLRENTSPRANKELDEAYVAVGSP 780
RKVKVLGSAFGTVYKGIWIPDGENKIPVAIVLRENTSPRANKELDEAYVAVGSP 780
QY 654 ----- 653
841 LVHRDLARNVLYKSPNHVKITDFGLARLIDIDETEHADGKVPKMMALSTLRRFT 900
LVHRDLARNVLYKSPNHVKITDFGLARLIDIDETEHADGKVPKMMALSTLRRFT 900
LVHRDLARNVLYKSPNHVKITDFGLARLIDIDETEHADGKVPKMMALSTLRRFT 900
QY 654 ----- 653
901 HQSDVSYGVTVWELMTFGAKPYDGIPIAREIPDLEKGERLPQPICTIDVYIMVYCM 960
HQSDVSYGVTVWELMTFGAKPYDGIPIAREIPDLEKGERLPQPICTIDVYIMVYCM 960
HQSDVSYGVTVWELMTFGAKPYDGIPIAREIPDLEKGERLPQPICTIDVYIMVYCM 960
QY 654 ----- 653
961 IDSECRREFRELVSERMAKRDQREVVYVIONEDLPASPLDSTFYRSLDDMDGLVDA 1020
IDSECRREFRELVSERMAKRDQREVVYVIONEDLPASPLDSTFYRSLDDMDGLVDA 1020
IDSECRREFRELVSERMAKRDQREVVYVIONEDLPASPLDSTFYRSLDDMDGLVDA 1020
QY 685 EETLVPOQGFCCPDPAAGAGMHHRR 712
EETLVPOQGFCCPDPAAGAGMHHRR 712
EETLVPOQGFCCPDPAAGAGMHHRR 712
QY 1021 EETLVPOQGFCCPDPAAGAGMHHRR 1048
EETLVPOQGFCCPDPAAGAGMHHRR 1048
EETLVPOQGFCCPDPAAGAGMHHRR 1048

RESULT 9
US-09-146-283-4
Sequence 4, Application US/09146283
Patent No. 5976546
GENERAL INFORMATION:
APPLICANT: Laus, Reiner
APPLICANT: Ruegg, Curtis L.
APPLICANT: Wu, Hongyu
TITLE OF INVENTION: Immunostimulatory Compositions
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Ave., Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/146, 283
FILING DATE: 03-SEPT-1998
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Judge, Linda R.

```

```

; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 782 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8
US-09-146-283-4

Query Match 91.9%; Score 3632; DB 2; Length 782;
Best Local Similarity 98.7%; Pred. No. 1.9e-305;
Matches 659; Conservative 0; Mismatches 5; Indels 4; Gaps 2;

QY 1 METALCRMGILLALLPPGAASTQVCTGTDMLRLPASPEHDMRLHYQCGVQVGNL 60
METALCRMGILLALLPPGAASTQVCTGTDMLRLPASPEHDMRLHYQCGVQVGNL 60
METALCRMGILLALLPPGAASTQVCTGTDMLRLPASPEHDMRLHYQCGVQVGNL 60
DQ 1 METALCRMGILLALLPPGAASTQVCTGTDMLRLPASPEHDMRLHYQCGVQVGNL 60
METALCRMGILLALLPPGAASTQVCTGTDMLRLPASPEHDMRLHYQCGVQVGNL 60
METALCRMGILLALLPPGAASTQVCTGTDMLRLPASPEHDMRLHYQCGVQVGNL 60
QY 61 ELTYLPNLSLFLQDIQEVGYVLIHNOVROVPLQRLRIYGTQLFEDNVALAVDNG 120
ELTYLPNLSLFLQDIQEVGYVLIHNOVROVPLQRLRIYGTQLFEDNVALAVDNG 120
ELTYLPNLSLFLQDIQEVGYVLIHNOVROVPLQRLRIYGTQLFEDNVALAVDNG 120
DQ 61 ELTYLPNLSLFLQDIQEVGYVLIHNOVROVPLQRLRIYGTQLFEDNVALAVDNG 120
ELTYLPNLSLFLQDIQEVGYVLIHNOVROVPLQRLRIYGTQLFEDNVALAVDNG 120
ELTYLPNLSLFLQDIQEVGYVLIHNOVROVPLQRLRIYGTQLFEDNVALAVDNG 120
QY 121 DPLNNTPTVYGASPGGRLQRLSTLEILKGVLIQRNPOLCYODIILMKDIFKNNOLA 180
DPLNNTPTVYGASPGGRLQRLSTLEILKGVLIQRNPOLCYODIILMKDIFKNNOLA 180
DPLNNTPTVYGASPGGRLQRLSTLEILKGVLIQRNPOLCYODIILMKDIFKNNOLA 180
DQ 121 DPLNNTPTVYGASPGGRLQRLSTLEILKGVLIQRNPOLCYODIILMKDIFKNNOLA 180
DPLNNTPTVYGASPGGRLQRLSTLEILKGVLIQRNPOLCYODIILMKDIFKNNOLA 180
DPLNNTPTVYGASPGGRLQRLSTLEILKGVLIQRNPOLCYODIILMKDIFKNNOLA 180
QY 181 LTLIDTNRSRACHPCSPMCKSGRGMWESSSEDCSLRTVCAGGCAKCKGLPTDCCHQC 240
LTLIDTNRSRACHPCSPMCKSGRGMWESSSEDCSLRTVCAGGCAKCKGLPTDCCHQC 240
LTLIDTNRSRACHPCSPMCKSGRGMWESSSEDCSLRTVCAGGCAKCKGLPTDCCHQC 240
DQ 181 LTLIDTNRSRACHPCSPMCKSGRGMWESSSEDCSLRTVCAGGCAKCKGLPTDCCHQC 240
LTLIDTNRSRACHPCSPMCKSGRGMWESSSEDCSLRTVCAGGCAKCKGLPTDCCHQC 240
LTLIDTNRSRACHPCSPMCKSGRGMWESSSEDCSLRTVCAGGCAKCKGLPTDCCHQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDFFSMNPDEGRYTFGASCTYAC 300
AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDFFSMNPDEGRYTFGASCTYAC 300
AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDFFSMNPDEGRYTFGASCTYAC 300
DQ 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDFFSMNPDEGRYTFGASCTYAC 300
AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDFFSMNPDEGRYTFGASCTYAC 300
AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDFFSMNPDEGRYTFGASCTYAC 300
QY 301 YNTLSTDVGSCTLVCPPLHNOEVTAEEDGTORCEKSKPCARVCYGLGMEHLREYRAVTSAN 360
YNTLSTDVGSCTLVCPPLHNOEVTAEEDGTORCEKSKPCARVCYGLGMEHLREYRAVTSAN 360
YNTLSTDVGSCTLVCPPLHNOEVTAEEDGTORCEKSKPCARVCYGLGMEHLREYRAVTSAN 360
DQ 301 YNTLSTDVGSCTLVCPPLHNOEVTAEEDGTORCEKSKPCARVCYGLGMEHLREYRAVTSAN 360
YNTLSTDVGSCTLVCPPLHNOEVTAEEDGTORCEKSKPCARVCYGLGMEHLREYRAVTSAN 360
YNTLSTDVGSCTLVCPPLHNOEVTAEEDGTORCEKSKPCARVCYGLGMEHLREYRAVTSAN 360
QY 361 IOEFAGCKITFGSLAFIPESFDGDPASNTAPLQPEQLQVETLEETGYLYISAMPDSL 420
IOEFAGCKITFGSLAFIPESFDGDPASNTAPLQPEQLQVETLEETGYLYISAMPDSL 420
IOEFAGCKITFGSLAFIPESFDGDPASNTAPLQPEQLQVETLEETGYLYISAMPDSL 420
DQ 361 IOEFAGCKITFGSLAFIPESFDGDPASNTAPLQPEQLQVETLEETGYLYISAMPDSL 420
IOEFAGCKITFGSLAFIPESFDGDPASNTAPLQPEQLQVETLEETGYLYISAMPDSL 420
IOEFAGCKITFGSLAFIPESFDGDPASNTAPLQPEQLQVETLEETGYLYISAMPDSL 420
QY 421 DLSVFQNLQVIRGRILHNGAYSLTLOGLSIWLGLSLRLSGSLALIHNTHLCEVHTV 480
DLSVFQNLQVIRGRILHNGAYSLTLOGLSIWLGLSLRLSGSLALIHNTHLCEVHTV 480
DLSVFQNLQVIRGRILHNGAYSLTLOGLSIWLGLSLRLSGSLALIHNTHLCEVHTV 480
DQ 421 DLSVFQNLQVIRGRILHNGAYSLTLOGLSIWLGLSLRLSGSLALIHNTHLCEVHTV 480
DLSVFQNLQVIRGRILHNGAYSLTLOGLSIWLGLSLRLSGSLALIHNTHLCEVHTV 480
DLSVFQNLQVIRGRILHNGAYSLTLOGLSIWLGLSLRLSGSLALIHNTHLCEVHTV 480
QY 481 PMDQLEFRNHQALHTANRDECEVGEGLACHOLCARHGHWGPPQCVNCSQFLRGQC 540
PMDQLEFRNHQALHTANRDECEVGEGLACHOLCARHGHWGPPQCVNCSQFLRGQC 540
PMDQLEFRNHQALHTANRDECEVGEGLACHOLCARHGHWGPPQCVNCSQFLRGQC 540
DQ 481 PMDQLEFRNHQALHTANRDECEVGEGLACHOLCARHGHWGPPQCVNCSQFLRGQC 540
PMDQLEFRNHQALHTANRDECEVGEGLACHOLCARHGHWGPPQCVNCSQFLRGQC 540
PMDQLEFRNHQALHTANRDECEVGEGLACHOLCARHGHWGPPQCVNCSQFLRGQC 540
QY 541 VEECRVLOGLPREYVNAHRLCPHPECOPQNSVTCGPEADOCVACAHKDDPPFCVAC 600
VEECRVLOGLPREYVNAHRLCPHPECOPQNSVTCGPEADOCVACAHKDDPPFCVAC 600
VEECRVLOGLPREYVNAHRLCPHPECOPQNSVTCGPEADOCVACAHKDDPPFCVAC 600
DQ 541 VEECRVLOGLPREYVNAHRLCPHPECOPQNSVTCGPEADOCVACAHKDDPPFCVAC 600
VEECRVLOGLPREYVNAHRLCPHPECOPQNSVTCGPEADOCVACAHKDDPPFCVAC 600
VEECRVLOGLPREYVNAHRLCPHPECOPQNSVTCGPEADOCVACAHKDDPPFCVAC 600
QY 601 PSGVKPDLSTYMPIMKFPDEGACQPCPINCTHSCVDLDKGCRAEORASPLTISVAVG 660
PSGVKPDLSYMPIMKFPDEGACQPCPINCTHSCVDLDKGCRAEORASPLTISVAVG 660
PSGVKPDLSYMPIMKFPDEGACQPCPINCTHSCVDLDKGCRAEORASPLTISVAVG 660
DQ 601 PSGVKPDLSTYMPIMKFPDEGACQPCPINCTHSCVDLDKGCRAEORASPLTISVAVG 660
PSGVKPDLSYMPIMKFPDEGACQPCPINCTHSCVDLDKGCRAEORASPLTISVAVG 660
PSGVKPDLSYMPIMKFPDEGACQPCPINCTHSCVDLDKGCRAEORASPLTISVAVG 660
QY 661 A-SPLDST 667
A-SPLDST 667
A-SPLDST 667
DQ 661 A-SPLDST 667
A-SPLDST 667
A-SPLDST 667
QY 658 ARSPSPST 665
ARSPSPST 665
ARSPSPST 665
DQ 658 ARSPSPST 665
ARSPSPST 665
ARSPSPST 665

RESULT 10
US-08-579-823A-4
Sequence 4, Application US/08579823A

```

```
/ Patent No. 6080409
/ GENERAL INFORMATION:
/ APPLICANT: Laus, Reiner
/ APPLICANT: Ruegg, Curtis L.
/ APPLICANT: Wu, Hongyu
/ TITLE OF INVENTION: Immunostimulatory Composition and Method
/ NUMBER OF SEQUENCES: 10
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Dehlinger & Associates
/ STREET: 350 Cambridge Ave. Suite 250
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94306
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/579,823A
/ FILING DATE: 03-DEC-1998
/ CLASSIFICATION: 536
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Judge, Linda R.
/ REGISTRATION NUMBER: 42,702
/ REFERENCE/DOCKET NUMBER: 7636-0010
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 650-324-0960
/ TELEFAX: 650-324-0960
/ INFORMATION FOR SEQ ID NO: 4:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 782 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHEICAL: NO
/ ORIGINAL SOURCE:
/ ORGANISM: homo sapiens
/ INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8
/ US-08-579-823A-4

Query Match          91.9%; Score 3632; DB 3; Length 782;
Best Local Similarity 98.7%; Pred. No. 1.9e-305;
Matches 659; Conservative 0; Mismatches 5; Indels 4; Gaps 2;
```

QY 1 MELAALCRWGLLLALPPGAASSTOYCTGTDKRLPASPETHLDMLRHLXOGCCYVQGNL 60  
DB 1 MELAALCRWGLLLALPPGAASSTOYCTGTDKRLPASPETHLDMLRHLXOGCCYVQGNL 60  
QY 61 ELTYIPTNASTSLFODIDYEVGYVIAHNOYRQVPLQRLRYRGQLFEDNYALAVLDNG 120  
DB 61 ELTYIPTNASTSLFODIDYEVGYVIAHNOYRQVPLQRLRYRGQLFEDNYALAVLDNG 120  
QY 121 DPLNNTTAVTASPGGLRELQRLSTETILKGVLLFORNPOLCYODTIIMKIDFHKNNOLA 180  
DB 121 DPLNNTTAVTASPGGLRELQRLSTETILKGVLLFORNPOLCYODTIIMKIDFHKNNOLA 180  
QY 181 LTLIDTNRACHPKSPCKSRCKGSESSDQCSTLRTVCAGGACRCKPLPTDCCHEQC 240  
DB 181 LTLIDTNRACHPKSPCKSRCKGSESSDQCSTLRTVCAGGACRCKPLPTDCCHEQC 240  
QY 241 AAGCGPRKHSQCLAHNHSIGICELHPALVTYNTDFFESMPNPEGRTYTGASCVTACP 300  
DB 241 AAGCGPRKHSQCLAHNHSIGICELHPALVTYNTDFFESMPNPEGRTYTGASCVTACP 300  
QY 301 YNYLSTDVGSCTIVCPILNQETAEADGTQRCCKSPCARVCYGLGMHLEVRVATGAN 360  
DB 301 YNYLSTDVGSCTIVCPILNQETAEADGTQRCCKSPCARVCYGLGMHLEVRVATGAN 360  
QY 361 IOEFGAGCKKRTIGSLAFLESTFDGDPASNTAPLOEQLOVFEETLEITGYLYISAMPD 420  
DB 361 IOEFGAGCKKRTIGSLAFLESTFDGDPASNTAPLOEQLOVFEETLEITGYLYISAMPD 420

```
QY 421 DLSTVFONLOVIRGRILHNGAYSLTLOGGLISWLGRLSELGSSGLALIHNTLCEVHTV 480  
DB 421 DLSTVFONLOVIRGRILHNGAYSLTLOGGLISWLGRLSELGSSGLALIHNTLCEVHTV 480  
QY 481 PWDOLFERNHQAALLHTANRPEDECVEGELACHQLCARGHCWGPPTQCVCNSQFLRGQEC 540  
DB 481 PWDOLFERNHQAALLHTANRPEDECVEGELACHQLCARGHCWGPPTQCVCNSQFLRGQEC 540  
QY 541 VEECRVLOGLPREYVNAHRCLEPCHECOPQNSVTCFGEADQCYACAHYKDPFCVARC 600  
DB 541 VEECRVLOGLPREYVNAHRCLEPCHECOPQNSVTCFGEADQCYACAHYKDPFCVARC 600  
QY 601 PSYVAPDLSTYMPIMKPEDEGACOPCPTNCTHSCVDLDKGPAPORASPLTSQMEDJGP 660  
DB 601 PSYVAPDLSTYMPIMKPEDEGACOPCPTNCTHSCVDLDKGPAPORASPLTSLE---AP 657  
QY 661 A-SPLDST 667  
DB 658 ARSPSPST 665
```

RESULT 11  
US-09-344-195-4  
Sequence 4, Application US/09344195  
Patent No. 6210662  
GENERAL INFORMATION:  
APPLICANT: Laus, Reiner  
Ruegg, Curtis L.  
Wu, Hongyu  
TITLE OF INVENTION: Immunostimulatory Compositions  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: 350 Cambridge Ave. Suite 250  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/344,195  
FILING DATE: 24-Jun-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/146,283  
FILING DATE: 03-SEPT-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Judge, Linda R.  
REGISTRATION NUMBER: 42,702  
REFERENCE/DOCKET NUMBER: 7636-0010.21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-324-0960  
TELEFAX: 650-324-0960  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 782 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: homo sapiens  
INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-344-195-4

Query Match 91.9%; Score 3632; DB 4; Length 782;  
Best Local Similarity 98.7%; Pred. No. 1.9e-305;





Db 601 ACQPCPINCTHSCVLDLDDKGPAAE 624

## RESULT 13

US-08-422-734-1

Sequence 1, Application US/08422734

Patent No. 6333169

GENERAL INFORMATION:

APPLICANT: Huddiak, Robert M.

APPLICANT: Shepard, H. Michael

TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/422,734

FILING DATE: 08/422,734

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/422108

FILING DATE: 14-Apr-1995

APPLICATION NUMBER: 08/355460

FILING DATE: 13-DEC-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/048346

FILING DATE: 15-APR-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/354319

FILING DATE: 19-MAY-1989

ATTORNEY/AGENT INFORMATION:

NAME: Lee, Wendy M

REGISTRATION NUMBER: 00,000

REFERENCE/DOCKET NUMBER: 554C2D1

TELEPHONE: 415/225-1994

TELEFAX: 415/952-9881

TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 624 amino acids

TYPE: Amino Acid

TOPOLOGY: Linear

US-08-422-734-1

Query Match 87.8% Score 3473; DB 4: Length 624;

Best Local Similarity 99.8% Pred. No. 7.7e-292;

Matches 623: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 22 STQVCTGDMKRLPASETHLDMRLHLYOGCOVGNLELTYLPTNLSLFDIOEYO 81  
1 STQVCTGDMKRLPASETHLDMRLHLYOGCOVGNLELTYLPTNLSLFDIOEYO 60  
Db 82 GYVLAHNOVROVPLQRLRIYVGTQLFEDNVALAVLDNGDPLNNTPTVTGASPGGLRELQ 141  
61 GYVLAHNOVROVPLQRLRIYVGTQLFEDNVALAVLDNGDPLNNTPTVTGASPGGLRELQ 120  
Db 142 LRLSLLEIKGGLVLIORNPOLCYODITLTKDIFHKNNOLATLIDNRSRACHPCSPMCKG 201  
121 LRLSLLEIKGGLVLIORNPOLCYODITLTKDIFHKNNOLATLIDNRSRACHPCSPMCKG 180  
Db 202 SRMGESSEDCOSLTRITVCAGGACARCKGPLPTDCHEOCAGCTGPKHSDCIACLHFNHS 261

Db 181 SRMGESSEDCOSLTRITVCAGGACARCKGPLPTDCHEOCAGCTGPKHSDCIACLHFNHS 240

Db 262 GICEHCPALVYNTDFESMPNPGRYTFGASCVYACPYNTLSTDVGSCTLVCPLANOE 321

Db 241 GICEHCPALVYNTDFESMPNPGRYTFGASCVYACPYNTLSTDVGSCTLVCPLANOE 300

Db 322 VTAEDGTORCEKSKPCARVCGIGLMEHLREVRATVSAIOEFACKKIFGSLAPLPSF 381

Db 301 VTAEDGTORCEKSKPCARVCGIGLMEHLREVRATVSAIOEFACKKIFGSLAPLPSF 360

Db 382 DGDPSANTAPLQPEOLQVETLEETGYLISAMPDLSPLDSVFNQYIRIRILHNGAY 441

Db 361 DGDPSANTAPLQPEOLQVETLEETGYLISAMPDLSPLDSVFNQYIRIRILHNGAY 420

Db 442 SLTLOGLSIWGLRSRLRLSGSLALIHNNHLCFVHTVPMQDLFRNPQALLHTRANPE 501

Db 421 SLTLOGLSIWGLRSRLRLSGSLALIHNNHLCFVHTVPMQDLFRNPQALLHTRANPE 480

Db 502 DECVGEGLACHQLCARGHMGPGPPQCVNCSQFLRGQECVEECRYLQGLPREYVARHCL 561

Db 481 DECVGEGLACHQLCARGHMGPGPPQCVNCSQFLRGQECVEECRYLQGLPREYVARHCL 540

Db 562 PCHPCEQONGSVTCFGEADQVACAHYKDPFCVARGPSGVKPDLSYMPIMKPPDEEG 621

Db 541 PCHPCEQONGSVTCFGEADQVACAHYKDPFCVARGPSGVKPDLSYMPIMKPPDEEG 600

Db 622 ACQPCPINCTHSCVLDLDDKGPAAE 645

Db 601 ACQPCPINCTHSCVLDLDDKGPAAE 624

## RESULT 14

US-09-630-155-2

Sequence 2, Application US/09630155

Patent No. 6414130

GENERAL INFORMATION:

APPLICANT: Doherty, Joni Kristin and Gail M. Clinton

TITLE OF INVENTION: HER-2 BINDING ANTAGONISTS

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: DAVIS WRIGHT TREMAINE LLP

STREET: 1501 Fourth Avenue, 2600 Century Square

CITY: Seattle

STATE: Washington

COUNTRY: U.S.A.

ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: PC compatible

OPERATING SYSTEM: Windows95

SOFTWARE: Word

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/630,155

FILING DATE: 16-Jan-2001

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Davison, Barry L.

REGISTRATION NUMBER: 47,309

REFERENCE/DOCKET NUMBER: 49321-10

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206 628-7621

TELEFAX: 206 628-7699

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 419

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: unknown

MOLECULE TYPE: polypeptide

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-630-155-2

```

Query Match      47.5%; Score 1878; DB 4; Length 419;
Best Local Similarity 83.0%; Pred. No. 3.3e-154;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

QY 1 METALCRGGLLALIPGASSTOCTGDMKRLPASPETHLMDLRHLVYOGCOVQVGNL 60
DB 1 METALCRGGLLALIPGASSTOCTGDMKRLPASPETHLMDLRHLVYOGCOVQVGNL 60
QY 61 ELTYLPNLSLFLDIQEOVGYLLAHNQROVPLQRLRIVRGTOLEFEDNVALVDNG 120
DB 61 ELTYLPNLSLFLDIQEOVGYLLAHNQROVPLQRLRIVRGTOLEFEDNVALVDNG 120
QY 121 DPLNNTPTVGTASPGGLRELQRLSLTEILKGVLIQNPOLCYDITLMKDIFKNNQLA 180
DB 121 DPLNNTPTVGTASPGGLRELQRLSLTEILKGVLIQNPOLCYDITLMKDIFKNNQLA 180
QY 181 LTLIDNRSRACHPCSPMCKSGRCKWESSSDQSLTRTVAGGACARCKGLPDDCCEOC 240
DB 181 LTLIDNRSRACHPCSPMCKSGRCKWESSSDQSLTRTVAGGACARCKGLPDDCCEOC 240
QY 241 AAGCTGPKHSDCLACHFNHSGICEHLCPALVYNTDTEFSMNPREGRYFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACHFNHSGICEHLCPALVYNTDTEFSMNPREGRYFGASCVTACP 300
QY 301 YNLSLTVGSCSTLVCPDLHNOEVTAEQDQRCCKSPCARVYGLGMEHLREYRAVTSAN 360
DB 301 YNLSLTVGSCSTLVCPDLHNOEVTAEQDQRCCKSPCARVYGLGMEHLREYRAVTSAN 360
QY 361 IOEFACKCKTFGSLAPLEPESFGDDPASNT---PPLQEOQLVFETLEITGYLISAMPD 417
DB 361 LRRQPG--PAHPVLSFLRPSWDLVSAFYSLPLAPLSTSVPI-----SPVSVGRGPD 405
QY 418 SLPLDSVFONLQVIRG 433
DB 406 --PDAAVAVNLSTRYEG 419

RESULT 15
US-08-484-438-7
; Sequence 7, Application US/08484438
; Patent No. 5811098
; Patent No. 5811098 5780031
; GENERAL INFORMATION:
; APPLICANT: Plowman, Gregory D.
; APPLICANT: Cuiusconu, Jean-Michel
; APPLICANT: Shoyab, Mohammed
; APPLICANT: Siegall, Clay B.
; APPLICANT: Hellstet m, Ingegerd
; TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,438
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/323,442
; FILING DATE: 14-OCT-1994
; APPLICATION NUMBER: US 08/150,704
; FILING DATE: 10-NOV-1993
; CLASSIFICATION: 530

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/981,165
; FILING DATE: 24-NOV-1992
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 5624-230
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1210 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-484-438-7

Query Match      38.8%; Score 1534; DB 2; Length 1210;
Best Local Similarity 42.9%; Pred. No. 7.7e-124;
Matches 283; Conservative 109; Mismatches 240; Indels 28; Gaps 6;

QY 11 LILALLPFGAA--STOYCTGDMKRLPASPETHLMDLRHLVYOGCOVQVGNLELTYLPN 68
DB 14 LILALCPASRALTEKKYCOQTSNKLTLQLEFEDHFLSLQRMFNCEVYGLNLEITVQRN 73
QY 69 ASLSFLDIQEOVGYLLAHNQROVPLQRLRIVRGTOLEFEDNVALVDNGDPLNNTPT 128
DB 74 YDLSFLKLTIOEVAGYLLAHNTVERIPLENLQIRGMNTEYALVALVSNYD----- 126
QY 129 VTGASPGGLRELQRLSLTEILKGVLIQNPOLCYDITLMKDIFKNNQLAFLTLIDTNR 188
DB 127 ---ANKTGLKELPMRNQELHGAVRSSNNPALCNESIGMWRIVSSDLSNNSMFOQN 183
QY 189 SRACHPCSPMCKSGRCKWESSSDQSLTRTVAGGCA--RCKGLPDDCCEOCAGCTCP 247
DB 184 LGSCQKDPSCPNVSGWAGAEENCOKLTKLCAQCSGRCKSPSCCHNOCAGCTCP 243
QY 248 KHSDCIACHFNHSGICEHLCPALVYNTDTEFSMNPREGRYFGASCVTACPYNLSTD 307
DB 244 RESDCLVCRKFRDCACTCKDCCPLMLNPTTYQMDVNPESKYSFGATCAKCRNTVVD 303
QY 308 VGSCTLVCPDLHNOEVTAEQDQRCCKSPCARVYGLGMEHLREYRAVTSANIOEFAC 367
DB 304 HGSCVRAACGADSYEM--BEDGVKCKCKCEGPCRVKVCNIGIGERKDSLSINATIKHKNC 362
QY 368 KTIFGSLAPLEPESFGDDPASNTAPLQEOQLVFETLEITGYLISAMPDLSLVFON 427
DB 363 TSIISGDHLILPVAFRGDSFTHTPLDPELDILKTYEITGFLIQAMPENRDLAHFEN 422
QY 428 LQYIRGILNAGVSLTLOGLISMLGLRSLRELGSLLALHNTHLCFVHTVPMQDLR 487
DB 423 LEIRGRTKHGOFSIAVLSNLTSTGLRSLKLEISDDVLTISGKNLCANTNMKKLFG 482
QY 488 NPHQALLHTANRPDEDCVSGELACHOLCARGHMGPPRTQCVNCSOFLRGQECVEECRYL 547
DB 483 TSGQKTIINRGBNCSKALQYCHALCSPEGCMGPRRCVCSGRANNSRRECVSKCKLL 542
QY 548 QGLPREVYNAHCLPCHRECOPOGNSYTCFPEADQCVACAHYKDPFCVAKRPSGVKPD 607
DB 543 EGEPRFEVENSEICQCHRECLPQAMNITCTGRGDNICQCAHYIDGPHCVKTCPPAGVGE 602
QY 608 LSTYPIKFPDEBACQPCPINCITGSCVDLDDKGCAPQASPLTISNEDLGASPLDST 667
DB 603 NNTL-VWKYADAGHYCHLCHPNCYTGCTGGLRECCPTN-----GPRISIAF 648

```

Search completed: January 13, 2003, 14:50:01  
Job time : 22.2232 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2003, 14:47:00 : Search time 8.2943 Seconds  
(without alignments)  
1665.428 Million cell updates/sec

Title: US-09-854-356-7  
Perfect score: 3954  
Sequence: 1 MELAICRMWGLLLALPPGA.....GFCPPDPAPGAGVHHRRH 712

Scoring table: BIOSUM62  
Gap 10.0 , Gapext 0.5

Searched: 118974 seqs, 19401057 residues

Total number of hits satisfying chosen parameters: 118974

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published\_Applications\_AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCR\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3954	100.0	712	9	US-09-854-356-7
2	3954	100.0	919	9	US-09-854-356-6
3	3776	95.5	1255	9	US-09-854-356-1
4	3776	95.5	1255	9	US-09-930-125-2
5	3776	95.5	1255	10	US-09-811-123-9
6	3776	95.5	1255	10	US-09-811-115-3
7	3727	94.3	1255	9	US-09-769-508-2
8	3628	91.8	653	9	US-09-854-356-3
9	3590	90.8	645	10	US-09-921-161-1
10	3209.5	81.2	1256	9	US-09-854-356-2
11	3205	81.1	1260	9	US-09-870-759-118
12	3189.5	80.7	1256	9	US-09-854-356-14
13	3110.5	78.7	654	9	US-09-854-356-8
14	1614.5	40.8	564	10	US-09-821-883-3
15	1614.5	40.8	697	10	US-09-821-883-4
16	1612	40.8	555	10	US-09-821-883-1
17	1612	40.8	690	10	US-09-821-883-2
18	1608	40.7	479	10	US-09-821-883-5
19	1587	40.1	289	10	US-09-821-883-23

20	1534	38.8	1210	10	US-09-725-433-2	Sequence 2, Appl1
21	1451	36.7	1308	10	US-09-940-101-2	Sequence 2, Appl1
22	1447	36.6	615	10	US-09-940-101-4	Sequence 4, Appl1
23	1028.5	26.0	478	10	US-09-867-521-2	Sequence 2, Appl1
24	350	8.9	583	9	US-09-930-125-9	Sequence 9, Appl1
25	350	8.9	587	9	US-09-930-125-8	Sequence 8, Appl1
26	350	8.9	589	9	US-09-930-125-10	Sequence 10, Appl1
27	350	8.9	600	9	US-09-930-125-11	Sequence 11, Appl1
28	326	8.2	59	9	US-09-854-356-5	Sequence 5, Appl1
29	326	8.2	266	9	US-09-854-356-4	Sequence 4, Appl1
30	311.5	7.9	1367	9	US-09-870-759-120	Sequence 120, App
31	291	7.4	370	10	US-09-205-658-104	Sequence 104, App
32	291	7.4	370	10	US-09-844-353A-104	Sequence 104, App
33	267.5	6.8	1724	10	US-09-205-658-112	Sequence 12, Appl
34	267.5	6.8	1724	10	US-09-844-353A-112	Sequence 12, Appl
35	250	6.3	366	10	US-09-205-658-103	Sequence 103, App
36	250	6.3	366	10	US-09-844-353A-103	Sequence 103, App
37	223.5	5.7	383	10	US-09-205-658-105	Sequence 105, App
38	223.5	5.7	383	10	US-09-844-353A-105	Sequence 105, App
39	212	5.4	381	10	US-09-205-658-106	Sequence 106, App
40	212	5.4	381	10	US-09-844-353A-106	Sequence 106, App
41	161	4.1	3084	10	US-09-938-275-4	Sequence 4, Appl1
42	160	4.0	1111	10	US-09-756-071B-15	Sequence 15, Appl1
43	160	4.0	1193	10	US-09-756-071B-13	Sequence 13, Appl1
44	159	4.0	3635	10	US-09-845-583-2	Sequence 2, Appl1
45	154	3.9	830	9	US-09-870-759-134	Sequence 134, App

#### ALIGNMENTS

RESULT 1  
US-09-854-356-7  
Sequence 7, Application US/09854356  
Patent No. US20020177567A1  
GENERAL INFORMATION:  
APPLICANT: Cheever, Martin A.  
APPLICANT: Chayesen, Dirk  
APPLICANT: Corixa Corporation  
APPLICANT: SmithKline Beecham Biologicals S. A.  
TITLE OF INVENTION: HER-2/neu Fusion Proteins  
FILE REFERENCE: 014058-009810PC  
CURRENT APPLICATION NUMBER: US/09/854,356  
CURRENT FILING DATE: 2001-05-09  
PRIOR APPLICATION NUMBER: US 09/493,480  
PRIOR FILING DATE: 2000-01-28  
PRIOR APPLICATION NUMBER: US 60/117,976  
PRIOR FILING DATE: 1999-01-29  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 7  
LENGTH: 712  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: fusion protein  
OTHER INFORMATION: of ECD and delta PD of human HER-2/neu  
US-09-854-356-7  
Query Match 100.0%; Score 3954; DB 9; Length 712;  
Best Local Similarity 100.0%; Pred. No. 1.6e-271;  
Matches 712; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MELAICRMWGLLLALPPGAASVCTGTDMKRLPASPETHLDMLRHLYGCGVVGNTL 60  
DB 1 MELAICRMWGLLLALPPGAASVCTGTDMKRLPASPETHLDMLRHLYGCGVVGNTL 60  
QY 61 ELTYPTNASSFLDDIDYGVGVIAHNOVRQVPLQRLRIYRGQLTEEDNATAVLVDNG 120  
DB 61 ELTYPTNASSFLDDIDYGVGVIAHNOVRQVPLQRLRIYRGQLTEEDNATAVLVDNG 120  
QY 121 DPLNNTFVTSGPSGLDELQRLSFTETIKGVLVLRORPDLCYODTILMKDIFHKNNOA 180  
DB 121 DPLNNTFVTSGPSGLDELQRLSFTETIKGVLVLRORPDLCYODTILMKDIFHKNNOA 180

Db	121	DP LNNTTPVVTGASPGGRLRELDLRSLTELEIKGVLIDQNPNQCYQDITLMDIDIFKNNQLA	189
Qy	181	LTLIDTNNSRACRHCSPMKCKSRCKWGESSEDCSLTTRVCAGGCAKCKGRLPTDCCHEOC	240
Db	181	LTLIDTNNSRACRHCSPMKCKSRCKWGESSEDCSLTTRVCAGGCAKCKGRLPTDCCHEOC	240
Qy	241	AAGCTGPRHSCCLACIHFNNHSGICEELHCPALVYTNTPTFESMPPEGRYTFGASCTYACP	3000
Db	241	AAGCTGPRHSCCLACIHFNNHSGICEELHCPALVYTNTPTFESMPPEGRYTFGASCTYACP	3000
Qy	301	YNYLSTDVGSCTLYCPLHNDQVTAEDGTQRCCKSKPCARCYGLGMEHLREYRAVYTSAN	3606
Db	301	YNYLSTDVGSCTLYCPLHNDQVTAEDGTQRCCKSKPCARCYGLGMEHLREYRAVYTSAN	3606
Qy	361	IOEFAGCKKIGSLAFLPESPDGSPASNTAPLOPEQLOVEFTLEETIGYLYISAMPDLSL	4202
Db	361	IOEFAGCKKIGSLAFLPESPDGSPASNTAPLOPEQLOVEFTLEETIGYLYISAMPDLSL	4202
Qy	421	DLVSQNIQVIRGRILHNGAVALTLQIGISLWLSRLSRLRETSGLALIHNTHTLCFEVHTV	4808
Db	421	DLVSQNIQVIRGRILHNGAVALTLQIGISLWLSRLSRLRETSGLALIHNTHTLCFEVHTV	4808
Qy	481	PMDOLFRRPHQALLHTANRPDEBCEVGBGLACHQLCANGHCWGEPITQCVNCSQFLRGQEC	5404
Db	481	PMDOLFRRPHQALLHTANRPDEBCEVGBGLACHQLCANGHCWGEPITQCVNCSQFLRGQEC	5404
Qy	541	VEECVVLQGLGREVYNAHNCILCPRECOPOKNGSVTCGFPREADQCVAAAHKKDPFECVARC	6000
Db	541	VEECVVLQGLGREVYNAHNCILCPRECOPOKNGSVTCGFPREADQCVAAAHKKDPFECVARC	6000
Qy	601	PSGVKPDLSYMPIMKFPDEBGACOPCPINCTHSCVDLDDKGCAPRQASPLTSONEDLGP	6606
Db	601	PSGVKPDLSYMPIMKFPDEBGACOPCPINCTHSCVDLDDKGCAPRQASPLTSONEDLGP	6606
Qy	661	ASPLDSTFYRSLDEDDDGDVLDAEYLYPQGGFCFCDPAPAGAGMHHNR	712
Db	661	ASPLDSTFYRSLDEDDDGDVLDAEYLYPQGGFCFCDPAPAGAGMHHNR	712
RESULT 2			
US-09-854-356-6			
Sequence 6, Application US/09854356			
Patent No. US2002017567A1			
GENERAL INFORMATION:			
APPLICANT: Cheever, Martin A.			
APPLICANT: Gheysen, Dirk			
APPLICANT: Corixa Corporation			
APPLICANT: SmithKline Beecham Biologicals S. A.			
TITLE OF INVENTION: HER-2/neu Fusion Proteins			
FILE REFERENCE: 014058-009810PC			
CURRENT APPLICATION NUMBER: US/09/854,356			
CURRENT FILING DATE: 2001-05-09			
PRIOR APPLICATION NUMBER: US 09/493,480			
PRIOR FILING DATE: 2000-01-28			
PRIOR APPLICATION NUMBER: US 60/117,976			
PRIOR FILING DATE: 1999-01-29			
NUMBER OF SEQ ID NOS: 26			
SOFTWARE: Patentin Ver. 2.1			
SEQ ID NO 6			
LENGTH: 919			
TYPE: PRT			
ORGANISM: Artificial Sequence			
FEATURE:			
OTHER INFORMATION: Description of Artificial Sequence: fusion protein			
US-09-854-356-6			

Query Match	100.0%	Score 3954;	DB 9;	Length 919;
Best Local Similarity	100.0%	Pred. No. 2.2e-271;		
Matches 712; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

D	b	1	MEALACWGLLALILPRGASQVQVGTDMKELRLPASPEHIDMLRHLYGQCVVGNL	60
O	y	61	ELTYLPTNASTSLFDIOIOVOGYVLIAHNOVROYVLORLRLVSTQLEFEDNYALAVDNG	120
D	b	61	ELTYLPTNASTSLFDIOIOVOGYVLIAHNOVROYVLORLRLVSTQLEFEDNYALAVDNG	120
O	y	121	DLINNTPTVTSAPGJBLELOLRSITLTELKGVLIORNPOLCYODTILMKDIFHKNOOLA	180
D	b	121	DLINNTPTVTSAPGJBLELOLRSITLTELKGVLIORNPOLCYODTILMKDIFHKNOOLA	180
O	y	181	LTLIDTNSRACHPCSPCKSGRCWGBESSBEDQSLRTVCAGGACARCKGRLPTDCCHEOC	240
D	b	181	LTLIDTNSRACHPCSPCKSGRCWGBESSBEDQSLRTVCAGGACARCKGRLPTDCCHEOC	240
O	y	241	AAGCGPRHSOCLACLIHNSGICELHCPALVYNTPTFESMPBERRYTFGASCVTACP	300
D	b	241	AAGCGPRHSOCLACLIHNSGICELHCPALVYNTPTFESMPBERRYTFGASCVTACP	300
O	y	301	YNYLSTDVSGSCTIACPLINQCEVTAEADGTORCEKCSKPCARCYGLGEMHILREVAVTAN	360
D	b	301	YNYLSTDVSGSCTIACPLINQCEVTAEADGTORCEKCSKPCARCYGLGEMHILREVAVTAN	360
O	y	361	IOERAGCKKIFGSLAFLESTFDGPRASNTAPLOPEOUVEFTLEETIGLYLISAMPDLP	420
D	b	361	IOERAGCKKIFGSLAFLESTFDGPRASNTAPLOPEOUVEFTLEETIGLYLISAMPDLP	420
O	y	421	DLSTVQNLQVIRGRILIHGAVSLTLOGIGLISWLBLSRLREJGSGIALIHNHTHLCFENVY	480
D	b	421	DLSTVQNLQVIRGRILIHGAVSLTLOGIGLISWLBLSRLREJGSGIALIHNHTHLCFENVY	480
O	y	481	PMDOLFRRPHOALLHTANRPDEBCVGBSLACHOLCARGHCKGPRPTQVWCISOFLRKQEC	540
D	b	481	PMDOLFRRPHOALLHTANRPDEBCVGBSLACHOLCARGHCKGPRPTQVWCISOFLRKQEC	540
O	y	541	VEECVVLQGLREYVYNAHNCILCPRECOPONGSUTCGRPEADQCVACAHKKDPFCVARC	600
D	b	541	VEECVVLQGLREYVYNAHNCILCPRECOPONGSUTCGRPEADQCVACAHKKDPFCVARC	600
O	y	601	PSGVKPDLSYMPIMKFPDEBAGACPRLINCTHSCVDLDDKGCAPAEORASPLTSONEDLGP	660
D	b	601	PSGVKPDLSYMPIMKFPDEBAGACPRLINCTHSCVDLDDKGCAPAEORASPLTSONEDLGP	660
O	y	661	ASPLDSTFYRSLBEDDDMGDLVDAEBEYLVPOQGFPCDDPAPGAGAMVHNHNR	712
D	b	661	ASPLDSTFYRSLBEDDDMGDLVDAEBEYLVPOQGFPCDDPAPGAGAMVHNHNR	712

```

RESULT 3
US-09-854-356-1
: Sequence 1, Application US/09854356
: Patent No. US20020177567A1
:
: GENERAL INFORMATION:
:
: APPLICANT: Cheever, Martin A.
: APPLICANT: Gheyssen, Dirk
: APPLICANT: Corixa Corporation
: APPLICANT: Smithline Becham Biologicals S. A.
: TITLE OF INVENTION: HER-2/neu Fusion Proteins
: FILE REFERENCE: 014058-009810PC
: CURRENT APPLICATION NUMBER: US/09/854, 356
: CURRENT FILING DATE: 2001-05-09
: PRIOR APPLICATION NUMBER: US 09/493, 480
: PRIOR FILING DATE: 2000-01-28
: PRIOR APPLICATION NUMBER: US 60/117, 976
: PRIOR FILING DATE: 1999-01-29
:
: NUMBER OF SEQ ID NOS: 26
:
: SOFTWARE: PatentIn Ver. 2.1
:
: SEQ ID NO 1
:
: LENGTH: 1255
:
: TYPE: PRT
:
: ORGANISM: Homo sapiens
:
: FEATURE:
:
: OTHER INFORMATION: human HER-2/neu protein
: NAME/KEY: DOMAIN

```

```
LOCATION: (1)..(653)
OTHER INFORMATION: extracellular domain (ECD)
NAME/KEY: DOMAIN
LOCATION: (676)..(1255)
OTHER INFORMATION: intracellular domain (ICD)
NAME/KEY: DOMAIN
LOCATION: (990)..(1255)
OTHER INFORMATION: phosphorylation domain (PD)
NAME/KEY: DOMAIN
LOCATION: (990)..(1048)
OTHER INFORMATION: fragment of the phosphorylation domain, preferred
OTHER INFORMATION: portion (delta PD)
US-09-854-356-1
```

```
Query Match          95.5%; Score 3776; DB 9; Length 1255;
Best Local Similarity 67.9%; Pred. No. 1.2e-258;
Matches 712; Conservative 0; Mismatches 0; Indels 336; Gaps 1;
```

```
OY 1 MELALACRGLLALLPPAASSTOYCTGDMKLRLPASPTHMDMLRHLTYGCGVYQGNL 60
DB 1 MELALACRGLLALLPPAASSTOYCTGDMKLRLPASPTHMDMLRHLTYGCGVYQGNL 60
OY 61 ELYLPTNALSFLDIOEVQGYVLIANNOVROVPLQRLIRIVGTQLFEDNYALAVLDNG 120
DB 61 ELYLPTNALSFLDIOEVQGYVLIANNOVROVPLQRLIRIVGTQLFEDNYALAVLDNG 120
OY 121 DPLNNTPTVTGASPGGLRLQRLSLTEILKGVLIQNRNQLCYODTILMKDIFHKNNQIA 180
DB 121 DPLNNTPTVTGASPGGLRLQRLSLTEILKGVLIQNRNQLCYODTILMKDIFHKNNQIA 180
OY 181 LFLIDNRSRACHPCSPMKSGSRGCESESDOSLTRVCAGAGCARCKPPLTDCCHQC 240
DB 181 LFLIDNRSRACHPCSPMKSGSRGCESESDOSLTRVCAGAGCARCKPPLTDCCHQC 240
OY 241 AAGCTGPKHSDCLACLFHFNHSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHFNHSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVTACP 300
OY 301 YNYLSTDVSGSCLVCPRLHNOEYTAEDGTQRCCKSPCARVCYGLGMEHLREVRATVTSAN 360
DB 301 YNYLSTDVSGSCLVCPRLHNOEYTAEDGTQRCCKSPCARVCYGLGMEHLREVRATVTSAN 360
OY 361 IOEPAGCKKIFGSLAFLPSPFDGDPASNTAPLOPQLOVFEETLEETGTYLISAMPDSL 420
DB 361 IOEPAGCKKIFGSLAFLPSPFDGDPASNTAPLOPQLOVFEETLEETGTYLISAMPDSL 420
OY 421 DLSVFQNLQVIRGRILHNGAVSLTLOGLGISWLGRLSRLRELSGLALIHNNHLCFVHTV 480
DB 421 DLSVFQNLQVIRGRILHNGAVSLTLOGLGISWLGRLSRLRELSGLALIHNNHLCFVHTV 480
OY 481 PWDQLFRNPHOALLHTANRPDECEYGEGLAQHQLCARHCGMGPPTQCYNSOFLRGQRC 540
DB 481 PWDQLFRNPHOALLHTANRPDECEYGEGLAQHQLCARHCGMGPPTQCYNSOFLRGQRC 540
OY 541 VECEKVLQGLPREYVNAHRCPCHPCCOPONGSVTCFGEAOCYACAHYKDPPECVANC 600
DB 541 VECEKVLQGLPREYVNAHRCPCHPCCOPONGSVTCFGEAOCYACAHYKDPPECVANC 600
OY 601 PSGVPRDLSYMPIMKFPPEDEGACQPCPINCSTHSCVDLDKGGPABGRASPISISAVYG 660
DB 601 PSGVPRDLSYMPIMKFPPEDEGACQPCPINCSTHSCVDLDKGGPABGRASPISISAVYG 660
OY 654 ----- 653
DB 661 ILVVVGLVFGILIRKROQIRKXYTMRRLQETELVEPLTPSGAMPNOAQRILKETEL 720
OY 654 ----- 653
DB 721 RKVKVLGSGAFGVYKGIWIPDENVKIPVAILKVLRENTSPANKETLDEAVYVAGVGP 780
OY 654 ----- 653
DB 781 YVSRLLIGLCTSTVOLVYQMLPYGCLLDHVRNRRGLSQDILLNMCMQIAKGMSTYLEDVR 840
```

```
OY 654 ----- 653
DB 841 LVHRDLAARNLVKSPNHWKTFDGLARLLDIDETERYHADGKVPILKMALESILRRPT 900
OY 654 ----- 653
DB 901 HOSDWSYGVYWEMLTPCAKPYDGIIPAREIDPLEKGERLPQPICTIDVYIMYKCM 960
OY 654 ----- 684
DB 961 IDSECRPRERELVEFSRMAPRQRFVYQNEIDLGASPLDSTFYRSLLEDMDGLVDA 1020
OY 685 EEYVPOGGEFCQPDPAAGAGVHHNRH 712
DB 1021 EEYVPOGGEFCQPDPAAGAGVHHNRH 1048
```

```
RESULT 4
US-09-930-125-2
Sequence 2, Application US/09930125
Publication No. US20020193329A1
GENERAL INFORMATION:
APPLICANT: Hand-Zimmerman, Susan
APPLICANT: Cheever, Martin A.
APPLICANT: Foy, Teresa M.
APPLICANT: Lodges, Michael J.
APPLICANT: Kalos, Michael D.
APPLICANT: McNeill, Patricia D.
APPLICANT: Vedicell, Thomas S.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
OF HER-2/NEU-ASSOCIATED MALIGNANCIES
FILE REFERENCE: 210121.544
CURRENT APPLICATION NUMBER: US/09/930,125
CURRENT FILING DATE: 2001-08-14
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows version 3.0
SEQ ID NO 2
LENGTH: 1255
TYPE: PRT
ORGANISM: Homo sapien
US-09-930-125-2
```

```
Query Match          95.5%; Score 3776; DB 9; Length 1255;
Best Local Similarity 67.9%; Pred. No. 1.2e-258;
Matches 712; Conservative 0; Mismatches 0; Indels 336; Gaps 1;
```

```
OY 1 MELALACRGLLALLPPAASSTOYCTGDMKLRLPASPTHMDMLRHLTYGCGVYQGNL 60
DB 1 MELALACRGLLALLPPAASSTOYCTGDMKLRLPASPTHMDMLRHLTYGCGVYQGNL 60
OY 61 ELYLPTNALSFLDIOEVQGYVLIANNOVROVPLQRLIRIVGTQLFEDNYALAVLDNG 120
DB 61 ELYLPTNALSFLDIOEVQGYVLIANNOVROVPLQRLIRIVGTQLFEDNYALAVLDNG 120
OY 121 DPLNNTPTVTGASPGGLRLQRLSLTEILKGVLIQNRNQLCYODTILMKDIFHKNNQIA 180
DB 121 DPLNNTPTVTGASPGGLRLQRLSLTEILKGVLIQNRNQLCYODTILMKDIFHKNNQIA 180
OY 181 LFLIDNRSRACHPCSPMKSGSRGCESESDOSLTRVCAGAGCARCKPPLTDCCHQC 240
DB 181 LFLIDNRSRACHPCSPMKSGSRGCESESDOSLTRVCAGAGCARCKPPLTDCCHQC 240
OY 241 AAGCTGPKHSDCLACLFHFNHSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHFNHSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVTACP 300
OY 301 YNYLSTDVSGSCLVCPRLHNOEYTAEDGTQRCCKSPCARVCYGLGMEHLREVRATVTSAN 360
DB 301 YNYLSTDVSGSCLVCPRLHNOEYTAEDGTQRCCKSPCARVCYGLGMEHLREVRATVTSAN 360
OY 361 IOEPAGCKKIFGSLAFLPSPFDGDPASNTAPLOPQLOVFEETLEETGTYLISAMPDSL 420
DB 361 IOEPAGCKKIFGSLAFLPSPFDGDPASNTAPLOPQLOVFEETLEETGTYLISAMPDSL 420
```

Db 361 IOEFAGCKITFSLAFIPESFDGDPASNTAPLOEOLQVETLEITGYLISAMPDLP 420  
QY 421 DLVSFQMLQVIRGRILNNGAVSLTLOGISWGLRSLRELSGLALIHNTLCEVHTV 480  
Db 421 DLVSFQMLQVIRGRILNNGAVSLTLOGISWGLRSLRELSGLALIHNTLCEVHTV 480  
QY 481 PMDLFRNPQALLHTANRPEDECVGEGSLACHOQACAGHCHGMPPTOCVNCISOFLRQEC 540  
Db 481 PMDLFRNPQALLHTANRPEDECVGEGSLACHOQACAGHCHGMPPTOCVNCISOFLRQEC 540  
QY 541 VEBECRVLOGLPREYVNAHCLPCHECOPONGSYTCGPEADQCVACAHKDPFCVARC 600  
Db 541 VEBECRVLOGLPREYVNAHCLPCHECOPONGSYTCGPEADQCVACAHKDPFCVARC 600  
QY 601 PSQVPLSLYMPIMKPFDEBEGACQPCPINCTHSCVDLDDKCPAERASPLTSLISAVYG 653  
Db 601 PSQVPLSLYMPIMKPFDEBEGACQPCPINCTHSCVDLDDKCPAERASPLTSLISAVYG 653  
QY 654 ----- 653  
Db 661 ILLVVLGVVFGILIKRROQKIRKYMRLLOETELVEPLTPSGAMPNOAMRLKETEL 720  
QY 654 ----- 653  
Db 721 RKVKVLGSGAFGVYKGIWIPDGENVKIPVAIKVIRENTSPKANKELIDEAYVAVGSP 780  
QY 654 ----- 653  
Db 781 YVSRLLGICLSTVQVLTQMLPFGCLLDHVRENHGRISQDLWMQOIAKMSYLEDR 840  
QY 654 ----- 653  
Db 841 LVHRDLAARNVLYKSPNHVKITDFGLARLDDIDETEHADGKVPKIMMALESILRRFT 900  
QY 654 ----- 653  
Db 901 HOSDVWSTGYVTWELMTFGAKPYGIPARETIDLEKGERLPDPCTIDVYIMTVKCM 960  
QY 654 ----- 653  
Db 961 IDSECRPRFRELVESEFSMARDPQRFVYIQNEDLGSPASPLDSTFYRSLLEDDMDGLVDA 1020  
QY 685 EEYLVPOOGFFCPDPAPGAGMHHRR 712  
Db 1021 EEYLVPOOGFFCPDPAPGAGMHHRR 1048

RESULT 5  
US-09-811-123-9  
Sequence 9, Application US/09811123  
Patent No. US2002001587A1  
GENERAL INFORMATION:  
APPLICANT: Sharon Erickson  
APPLICANT: Ralph Schwall  
APPLICANT: Mark Sliwowski  
TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-ETDB  
FILE REFERENCE: GENENT. 073A2  
CURRENT APPLICATION NUMBER: US/09/811,123  
CURRENT FILING DATE: 2001-03-16  
PRIOR APPLICATION NUMBER: 60/238,327  
PRIOR FILING DATE: 2000-10-05  
PRIOR APPLICATION NUMBER: 09/602,530  
PRIOR FILING DATE: 2000-06-23  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 9  
LENGTH: 1255  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-811-123-9

Query Match 95.5%; Score 3776; DB 10; Length 1255;

Best Local Similarity 67.9%; Pred. No. 1.2e-258;  
Matches 712; Conservative 0; Mismatches 0; Indels 336; Gaps 1;  
QY 1 MELAALCRWGLLALLPPGAASVQCTGTDKMLPASPETHLMDLHLYVQGVQVGNL 60  
Db 1 MELAALCRWGLLALLPPGAASVQCTGTDKMLPASPETHLMDLHLYVQGVQVGNL 60  
QY 61 ELTYLPTNASLSFQDIOEVQGYVLIHNGVQVPLRLRIVRSTQOLFEDNYALAVDNG 120  
Db 61 ELTYLPTNASLSFQDIOEVQGYVLIHNGVQVPLRLRIVRSTQOLFEDNYALAVDNG 120  
QY 121 DPLNNTPTVYGASPGGLREQLRSLTEILGKGVLIQRNPOLCYODTILMKDIFKNNOLA 180  
Db 121 DPLNNTPTVYGASPGGLREQLRSLTEILGKGVLIQRNPOLCYODTILMKDIFKNNOLA 180  
QY 181 LTLIDTNRSRACHPCSPCKGSRMGSSSEDCOSTLTFTVACGACRKGPLPTDCHQEC 240  
Db 181 LTLIDTNRSRACHPCSPCKGSRMGSSSEDCOSTLTFTVACGACRKGPLPTDCHQEC 240  
QY 241 AAGCTGPKHSDCLACLFHNHSGICELCPALVTYNTDTFESMPREGRYTFGASCYACP 300  
Db 241 AAGCTGPKHSDCLACLFHNHSGICELCPALVTYNTDTFESMPREGRYTFGASCYACP 300  
QY 301 YNYLSTVGSCTLVCPILHNDVTAEDGTORCEKSKPCARCYGLGMEHLREYRAVTSAN 360  
Db 301 YNYLSTVGSCTLVCPILHNDVTAEDGTORCEKSKPCARCYGLGMEHLREYRAVTSAN 360  
QY 361 IOEFAGCKITFSLAFIPESFDGDPASNTAPLOEOLQVETLEITGYLISAMPDLP 420  
Db 361 IOEFAGCKITFSLAFIPESFDGDPASNTAPLOEOLQVETLEITGYLISAMPDLP 420  
QY 421 DLVSFQMLQVIRGRILNNGAVSLTLOGISWGLRSLRELSGLALIHNTLCEVHTV 480  
Db 421 DLVSFQMLQVIRGRILNNGAVSLTLOGISWGLRSLRELSGLALIHNTLCEVHTV 480  
QY 481 PMDLFRNPQALLHTANRPEDECVGEGSLACHOQACAGHCHGMPPTOCVNCISOFLRQEC 540  
Db 481 PMDLFRNPQALLHTANRPEDECVGEGSLACHOQACAGHCHGMPPTOCVNCISOFLRQEC 540  
QY 541 VEBECRVLOGLPREYVNAHCLPCHECOPONGSYTCGPEADQCVACAHKDPFCVARC 600  
Db 541 VEBECRVLOGLPREYVNAHCLPCHECOPONGSYTCGPEADQCVACAHKDPFCVARC 600  
QY 601 PSQVPLSLYMPIMKPFDEBEGACQPCPINCTHSCVDLDDKCPAERASPLTSLISAVYG 653  
Db 601 PSQVPLSLYMPIMKPFDEBEGACQPCPINCTHSCVDLDDKCPAERASPLTSLISAVYG 653  
QY 654 ----- 653  
Db 661 ILLVVLGVVFGILIKRROQKIRKYMRLLOETELVEPLTPSGAMPNOAMRLKETEL 720  
QY 654 ----- 653  
Db 721 RKVKVLGSGAFGVYKGIWIPDGENVKIPVAIKVIRENTSPKANKELIDEAYVAVGSP 780  
QY 654 ----- 653  
Db 781 YVSRLLGICLSTVQVLTQMLPFGCLLDHVRENHGRISQDLWMQOIAKMSYLEDR 840  
QY 654 ----- 653  
Db 841 LVHRDLAARNVLYKSPNHVKITDFGLARLDDIDETEHADGKVPKIMMALESILRRFT 900  
QY 654 ----- 653  
Db 901 HOSDVWSTGYVTWELMTFGAKPYGIPARETIDLEKGERLPDPCTIDVYIMTVKCM 960  
QY 654 ----- 653  
Db 961 IDSECRPRFRELVESEFSMARDPQRFVYIQNEDLGSPASPLDSTFYRSLLEDDMDGLVDA 1020  
QY 685 EEYLVPOOGFFCPDPAPGAGMHHRR 712



DB 1021 EEYLVPOGFCPPDPAPGAGMVHHRH 1048

RESULT 6

US-09-811-115-3

Sequence 3, Application US/09811115

Patent No. US20020035736A1

GENERAL INFORMATION:

APPLICANT: Erickson, Sharon

APPLICANT: King, Kathleen

TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL

FILE REFERENCE: GENE 034A

CURRENT APPLICATION NUMBER: US/09/811,115

PRIOR FILING DATE: 2001-03-16

PRIOR APPLICATION NUMBER: 60/189,844

NUMBER OF SEQ ID NOS: 4

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 3

LENGTH: 1255

TYPE: PRT

ORGANISM: Homo sapiens

US-09-811-115-3

Query Match 95.5%; Score 3776; DB 10; Length 1255;

Best Local Similarity 67.9%; Pred. No. 1.2e-258;

Matches 712; Conservative 0; Mismatches 0; Indels 336; Gaps 1;

QY 1 MELALCRWGLLALLPPGAASVCTGTDMKRLPASPETHLDMRLHLYOGCQVVGNTL 60

DB 1 MELALCRWGLLALLPPGAASVCTGTDMKRLPASPETHLDMRLHLYOGCQVVGNTL 60

QY 61 ETTYPTNASSLFODIOEVGYVLIANOVROVPLQRLRIYRGQLEFEDNYALAVLNG 120

DB 61 ETTYPTNASSLFODIOEVGYVLIANOVROVPLQRLRIYRGQLEFEDNYALAVLNG 120

QY 121 DPLNNTPTVTGASPGGLRELDRLSTLTKGVLQORNPOLCYOPTILMKDIFHKNOLA 180

DB 121 DPLNNTPTVTGASPGGLRELDRLSTLTKGVLQORNPOLCYOPTILMKDIFHKNOLA 180

QY 181 LTLIDTNSRACHPCSPCKSGRCWGESSEDCOSLTRFYAGGACARCKGRLPTDCHEOC 240

DB 181 LTLIDTNSRACHPCSPCKSGRCWGESSEDCOSLTRFYAGGACARCKGRLPTDCHEOC 240

QY 241 AAGCTGPRHSDCLAHFNHSGICELHCPALVTYNTDTFESMPNEGRYTFGASCVTACP 300

DB 241 AAGCTGPRHSDCLAHFNHSGICELHCPALVTYNTDTFESMPNEGRYTFGASCVTACP 300

QY 301 YNYLSTDVGSCTVCPRLNQEYTAEDGTORCEKSKPCARVCYGLGMHRLREYRAVTSAN 360

DB 301 YNYLSTDVGSCTVCPRLNQEYTAEDGTORCEKSKPCARVCYGLGMHRLREYRAVTSAN 360

QY 361 IOEFAGCKKIGSLAFPESEFDGDPASNTAPLQPEOLQVFEITGELYYSANPDSL 420

DB 361 IOEFAGCKKIGSLAFPESEFDGDPASNTAPLQPEOLQVFEITGELYYSANPDSL 420

QY 421 DLSVFQNLQVIRGRILHNGAVSLTQGLIGISMLGRLSRLREGSLALIHNTHLCEFTV 480

DB 421 DLSVFQNLQVIRGRILHNGAVSLTQGLIGISMLGRLSRLREGSLALIHNTHLCEFTV 480

QY 481 PMDOLFRRPHALLHTARPEDECEGELACHOLCARHCWCPRTQCVNCSQFLRGEC 540

DB 481 PMDOLFRRPHALLHTARPEDECEGELACHOLCARHCWCPRTQCVNCSQFLRGEC 540

QY 541 VEECRVLOGLPREYVNAHCLPCHECOPONGSVTCFEPREADQCVACAHYKDPFCVARG 600

DB 541 VEECRVLOGLPREYVNAHCLPCHECOPONGSVTCFEPREADQCVACAHYKDPFCVARG 600

QY 601 PSQVAPDLSTYPIKFPDEBAGACOPRLNCTHSCVDLDDKCPAQARASPLTS----- 653

DB 601 PSQVAPDLSTYPIKFPDEBAGACOPRLNCTHSCVDLDDKCPAQARASPLTS----- 653

QY 601 PSQVAPDLSTYPIKFPDEBAGACOPRLNCTHSCVDLDDKCPAQARASPLTS----- 653

DB 601 PSQVAPDLSTYPIKFPDEBAGACOPRLNCTHSCVDLDDKCPAQARASPLTS----- 653

QY 654 ----- 653

DB 661 ILLVYVVGIVFGLIKRROKIRKTYMRLLOETELVEHLPSPGAMPQAOIRIKETEL 720

QY 654 ----- 653

DB 721 RAVVILSGAFGTYYKGIWIPDGENVKIPAIKVLRENTSPKANKEILDEAYVMAGVSP 780

QY 654 ----- 653

DB 781 YVSRLLGICLSTVOLVQVLMFYGLLDHVRNRRGLSQDILNMQOIAKMSYLEDVR 840

QY 654 ----- 653

DB 841 LVHRLAARNVLKSPNHVKTIDPGLARLLIDETEHADGKVPDKMALESLIRRET 900

QY 654 ----- 653

DB 901 HQSDVSYGVYVWELMTFGAKPYDIPAREIPDLLEKGERLPQPICITDYVMIMVKCM 960

QY 654 ----- 653

DB 961 IDSECRPRFRELVSFSESMARDPQRFVYIQNEDLGPASPLDSTFYRSLLEDMDGLVDA 1020

QY 685 EEYLVPOGFCPPDPAPGAGMVHHRH 712

DB 1021 EEYLVPOGFCPPDPAPGAGMVHHRH 1048

RESULT 7

US-09-769-508-2

Sequence 2, Application US/09769508

Patent No. US20020155527A1

GENERAL INFORMATION:

APPLICANT: STUART, SUSAN G.

APPLICANT: MONAHAN, JOHN J.

APPLICANT: LANGTON, BEATRICE CLAUDIA

APPLICANT: HANCOCK, MIRIAM E.C.

APPLICANT: CHAO, LORRINE A.

TITLE OF INVENTION: C-ERBB-2 EXTERNAL DOMAIN: GP75

FILE REFERENCE: BEBIO-111-C1

CURRENT APPLICATION NUMBER: US/09/769,508

CURRENT FILING DATE: 2001-01-26

NUMBER OF SEQ ID NOS: 2

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2

LENGTH: 1255

TYPE: PRT

ORGANISM: Homo sapiens

US-09-769-508-2

Query Match 94.3%; Score 3727; DB 9; Length 1255;

Best Local Similarity 67.5%; Pred. No. 3.3e-255;

Matches 707; Conservative 0; Mismatches 5; Indels 336; Gaps 1;

QY 1 MELALCRWGLLALLPPGAASVCTGTDMKRLPASPETHLDMRLHLYOGCQVVGNTL 60

DB 1 MELALCRWGLLALLPPGAASVCTGTDMKRLPASPETHLDMRLHLYOGCQVVGNTL 60

QY 61 ETTYPTNASSLFODIOEVGYVLIANOVROVPLQRLRIYRGQLEFEDNYALAVLNG 120

DB 61 ETTYPTNASSLFODIOEVGYVLIANOVROVPLQRLRIYRGQLEFEDNYALAVLNG 120

QY 121 DPLNNTPTVTGASPGGLRELDRLSTLTKGVLQORNPOLCYOPTILMKDIFHKNOLA 180

DB 121 DPLNNTPTVTGASPGGLRELDRLSTLTKGVLQORNPOLCYOPTILMKDIFHKNOLA 180

QY 181 LTLIDTNSRACHPCSPCKSGRCWGESSEDCOSLTRFYAGGACARCKGRLPTDCHEOC 240

DB 181 LTLIDTNSRACHPCSPCKSGRCWGESSEDCOSLTRFYAGGACARCKGRLPTDCHEOC 240

QY 241 AAGCTGPRHSDCLAHFNHSGICELHCPALVTYNTDTFESMPNEGRYTFGASCVTACP 300

```

|||||
Db 241 AACCTGKHSDDCLACLFHNSGICEHLCPALVYNTDTFESMPDEGRYTFGASCVTACP 300
QY 301 YNTLSTVGSCTLVCPHNEVTAEDGTORCEKSKPCAVCYGLGMEHLREVAVTAN 360
Db 301 YNTLSTVGSCTLVCPHNEVTAEDGTORCEKSKPCAVCYGLGMEHLREVAVTAN 360
QY 361 IOEFAGCKIFGSLAFIPESFDGDPASNTAPLOPEQLQVEETELETGYLYISAMPDLP 420
Db 361 IOEFAGCKIFGSLAFIPESFDGDPASNTAPLOPEQLQVEETELETGYLYISAMPDLP 420
QY 421 DLSVFQMLQYIRGRILHNGAYSLTLOGLSISWGLRSLRELGSGLAIHHNTHLCFVHTV 480
Db 421 DLSVFQMLQYIRGRILHNGAYSLTLOGLSISWGLRSLRELGSGLAIHHNTHLCFVHTV 480
QY 481 PMDOLFNPQALHTANREDECEVSGELACHOLCARGHGMGPPTQCVNCSOFLRGQEC 540
Db 481 PMDOLFNPQALHTANREDECEVSGELACHOLCARGHGMGPPTQCVNCSOFLRGQEC 540
QY 541 VEECRVLOGLPREVYNARHCLPCHPECOPONGSVTCFGRPADQCVACAHYKDPFCVARC 600
Db 541 VEECRVLOGLPREVYNARHCLPCHPECOPONGSVTCFGRPADQCVACAHYKDPFCVARC 600
QY 601 PSGVKPDLSTYMPIMKFPDEEGACQPCPINCTHSCVDLDDKGCRAEQASPLTS 653
Db 601 PSGVKPDLSTYMPIMKFPDEEGACQPCPINCTHSCVDLDDKGCRAEQASPLTS 653
QY 654 ----- 653
Db 661 ILVVVVGIVFGILIKRROOKIKKTYMRLLQETELVEPLTPSGAMPNOMRIKETEL 720
QY 654 ----- 653
Db 721 RKVKVLGSGAFGYKGIWIPDGENKIPVAIKVLRNTSPKANKETLDEAYVAVGSP 780
QY 654 ----- 653
Db 781 YSRLLGICLTSTYQVLTQMPYGCLLDHVRENKRLSODLLNMCQIAKGMSTLEDVR 840
QY 654 ----- 653
Db 841 LVHRDLAARVAVKSPNHVITDGLAKRLDIDETEHADGKVPITWMALESILRRFT 900
QY 654 ----- 653
Db 901 HOSDWSYGVYVWELMFGAKPYDIPARELPDLLEKGERLPORPCTIDYVIMVWCM 960
QY 654 ----- 960
Db 961 IDSECRPREBELVSEFSRMAADPQREVVYIQNEDLGPA SPLDSTFYRSLLDEDDMGDLVDA 1020
QY 685 EBYLVPOQGFCDPAPAGAGMHHNRH 712
Db 1021 EBYLVPOQGFCDPAPAGAGMHHNRH 1048

```

RESULT 8  
US-09-854-356-3  
; Sequence 3, Application US/09854356  
; Patent No. US20020177567A1

GENERAL INFORMATION:  
; APPLICANT: Cheever, Martin A.  
; APPLICANT: Cheever, Dirk  
; APPLICANT: Corixa Corporation  
; APPLICANT: SmithKline Beecham Biologicals S. A.  
; TITLE OF INVENTION: HER-2/neu Fusion Proteins  
; FILE REFERENCE: 014058-009810PC  
; CURRENT APPLICATION NUMBER: US/09/854,356  
; CURRENT FILING DATE: 2001-05-09  
; PRIOR APPLICATION NUMBER: US 09/493,480  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: US 60/117,976  
; PRIOR FILING DATE: 1999-01-29

```

; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 653
; TYPE: PR1
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: extracellular domain (ECD) of human HER-2/neu
US-09-854-356-3

```

```

Query Match          91.8%; Score 3628; DB 9; Length 653;
Best Local Similarity 100.0%; Pred. No. 1,4e-248;
Matches 653; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELAALCRWGLLALLPPGAASPOVCTGTDMKRLRPSPEHIDMLRHLVQGGVOVGNL 60
Db 1 MELAALCRWGLLALLPPGAASPOVCTGTDMKRLRPSPEHIDMLRHLVQGGVOVGNL 60
QY 61 ELTYLPTNASLSFLQDIQEVQYVLIANNOVROYPLQRLRIVNGTQLFEDNVALLVLDNG 120
Db 61 ELTYLPTNASLSFLQDIQEVQYVLIANNOVROYPLQRLRIVNGTQLFEDNVALLVLDNG 120
QY 121 DPLNNTPTVVGASPGGLREIQLRSLPTEILKGVLIQNPOLCYODTILMKDIFHKNNOLA 180
Db 121 DPLNNTPTVVGASPGGLREIQLRSLPTEILKGVLIQNPOLCYODTILMKDIFHKNNOLA 180
QY 181 LTLIDTNRSRACHPCSPMGKSGRCWGESSESDCSLRTVCAGCARGKGLPTDCCHEQC 240
Db 181 LTLIDTNRSRACHPCSPMGKSGRCWGESSESDCSLRTVCAGCARGKGLPTDCCHEQC 240
QY 241 AAGCTGKHSDDCLACLFHNSGICEHLCPALVYNTDTFESMPDEGRYTFGASCVTACP 300
Db 241 AAGCTGKHSDDCLACLFHNSGICEHLCPALVYNTDTFESMPDEGRYTFGASCVTACP 300
QY 301 YNTLSTVGSCTLVCPHNEVTAEDGTORCEKSKPCAVCYGLGMEHLREVAVTAN 360
Db 301 YNTLSTVGSCTLVCPHNEVTAEDGTORCEKSKPCAVCYGLGMEHLREVAVTAN 360
QY 361 IOEFAGCKIFGSLAFIPESFDGDPASNTAPLOPEQLQVEETELETGYLYISAMPDLP 420
Db 361 IOEFAGCKIFGSLAFIPESFDGDPASNTAPLOPEQLQVEETELETGYLYISAMPDLP 420
QY 421 DLSVFQMLQYIRGRILHNGAYSLTLOGLSISWGLRSLRELGSGLAIHHNTHLCFVHTV 480
Db 421 DLSVFQMLQYIRGRILHNGAYSLTLOGLSISWGLRSLRELGSGLAIHHNTHLCFVHTV 480
QY 481 PMDOLFNPQALHTANREDECEVSGELACHOLCARGHGMGPPTQCVNCSOFLRGQEC 540
Db 481 PMDOLFNPQALHTANREDECEVSGELACHOLCARGHGMGPPTQCVNCSOFLRGQEC 540
QY 541 VEECRVLOGLPREVYNARHCLPCHPECOPONGSVTCFGRPADQCVACAHYKDPFCVARC 600
Db 541 VEECRVLOGLPREVYNARHCLPCHPECOPONGSVTCFGRPADQCVACAHYKDPFCVARC 600
QY 601 PSGVKPDLSTYMPIMKFPDEEGACQPCPINCTHSCVDLDDKGCRAEQASPLTS 653
Db 601 PSGVKPDLSTYMPIMKFPDEEGACQPCPINCTHSCVDLDDKGCRAEQASPLTS 653

```

RESULT 9  
US-09-921-161-1  
; Sequence 1, Application US/09921161  
; Patent No. US20020090662A1  
GENERAL INFORMATION:  
; APPLICANT: Ralph, Peter  
; APPLICANT: ANALYTICAL METHOD  
; FILE REFERENCE: GENENT, 066A  
; CURRENT APPLICATION NUMBER: US/09/921,161  
; CURRENT FILING DATE: 2001-08-01  
; PRIOR APPLICATION NUMBER: 60/225,433  
; PRIOR FILING DATE: 2000-08-15  
; NUMBER OF SEQ ID NOS: 1  
; SOFTWARE: FastSeq for Windows Version 4.0

```

; SEQ ID NO 1
; LENGTH: 645
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-921-161-1

```

```

Query Match          90.8%; Score 3590; DB 10; Length 645;
Best Local Similarity 100.0%; Pred. No. 6.8e-246;
Matches 645; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 MELALCRNGLLALPPGAASVQVCTGDMKRLRASPETHDMRLHYOGCQVQGNL 60
DB 1 MELALCRNGLLALPPGAASVQVCTGDMKRLRASPETHDMRLHYOGCQVQGNL 60
OY 61 ELTYPTNASLSFLDIOEVQGVYLAHNOVQVPLQRLRIYRGTLDFEDNTALAVLNG 120
DB 61 ELTYPTNASLSFLDIOEVQGVYLAHNOVQVPLQRLRIYRGTLDFEDNTALAVLNG 120
OY 121 DPLNNTPTVTGASPGGLRELQRLSTLEILKGGVLIQIRNPOLCYQDTILMKDIFHKNNOLA 180
DB 121 DPLNNTPTVTGASPGGLRELQRLSTLEILKGGVLIQIRNPOLCYQDTILMKDIFHKNNOLA 180
OY 181 LTLIDTNRSRACHPCSPMKGSKRCWGESSEDDCSLRTFYCAGGACARCKGRPLPTDCCHQC 240
DB 181 LTLIDTNRSRACHPCSPMKGSKRCWGESSEDDCSLRTFYCAGGACARCKGRPLPTDCCHQC 240
OY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNREGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNREGRYTFGASCVTACP 300
OY 301 YNLTSDVGSCTLVCPHNOEYTAEDGTORCEKSKPCARVCYGLGMELREVRATYSAN 360
DB 301 YNLTSDVGSCTLVCPHNOEYTAEDGTORCEKSKPCARVCYGLGMELREVRATYSAN 360
OY 361 IOEFAGCKKIFGSLAFLEPSPFDGPNASNTAPLOPEQLQVFETLEETIGLYISAMPDSL 420
DB 361 IOEFAGCKKIFGSLAFLEPSPFDGPNASNTAPLOPEQLQVFETLEETIGLYISAMPDSL 420
OY 421 DLSVFQNLQVIRIGRIHNGAYSILTQGLISWLGRLSRLRELSGGLAIHHNTHLCFVHT 480
DB 421 DLSVFQNLQVIRIGRIHNGAYSILTQGLISWLGRLSRLRELSGGLAIHHNTHLCFVHT 480
OY 481 PMDOLFNRPHOALLHTANRPEDCEVGEGLACHQOLCARHGWGPPTQVNCISOFLSGQSC 540
DB 481 PMDOLFNRPHOALLHTANRPEDCEVGEGLACHQOLCARHGWGPPTQVNCISOFLSGQSC 540
OY 541 VEECRVLOGLPREYVNAHRCPCHEPCOPQNGSVTCFEPADQCVACAHYKDPPECVAC 600
DB 541 VEECRVLOGLPREYVNAHRCPCHEPCOPQNGSVTCFEPADQCVACAHYKDPPECVAC 600
OY 601 PGVVPDLSTYMPIMKFPDEEGACOPPCINCHTSCVDLDDKCGPAE 645
DB 601 PGVVPDLSTYMPIMKFPDEEGACOPPCINCHTSCVDLDDKCGPAE 645

```

```

RESULT 10
US-09-854-356-2

```

```

; Sequence 2, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheysen, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854.356
; PRIOR FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26

```

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1256
; TYPE: PRT
; ORGANISM: Rattus sp.
; FEATURE:
; OTHER INFORMATION: rat HER-2/neu protein
; NAME/KEY: DOMAIN
; LOCATION: (1)..(654)
; OTHER INFORMATION: extracellular domain (ECD)
; NAME/KEY: DOMAIN
; LOCATION: (677)..(1256)
; OTHER INFORMATION: intracellular domain (ICD)
; NAME/KEY: DOMAIN
; LOCATION: (721)..(998)
; OTHER INFORMATION: kinase domain (KD)
; NAME/KEY: DOMAIN
; LOCATION: (991)..(1049)
; OTHER INFORMATION: phosphorylation domain (PD)
; OTHER INFORMATION: fragment of the phosphorylation domain, preferred
; OTHER INFORMATION: portion (delta PD)
US-09-854-356-2

```

```

Query Match          81.2%; Score 3209.5; DB 9; Length 1256;
Best Local Similarity 58.0%; Pred. No. 1.1e-218;
Matches 608; Conservative 34; Mismatches 70; Indels 337; Gaps 2;

```

```

OY 1 MELALCRNGLLALPPGAASVQVCTGDMKRLRASPETHDMRLHYOGCQVQGNL 60
DB 1 MELALCRNGLLALPPGAASVQVCTGDMKRLRASPETHDMRLHYOGCQVQGNL 60
OY 61 ELTYPTNASLSFLDIOEVQGVYLAHNOVQVPLQRLRIYRGTLDFEDNTALAVLNG 120
DB 61 ELTYPTNASLSFLDIOEVQGVYLAHNOVQVPLQRLRIYRGTLDFEDNTALAVLNG 120
OY 121 DPLNNTPTVT-GASPGGLRELQRLSTLEILKGGVLIQIRNPOLCYQDTILMKDIFHKNNOL 179
DB 121 DPLNNTPTVTGASPGGLRELQRLSTLEILKGGVLIQIRNPOLCYQDTILMKDIFHKNNOL 180
OY 180 ALTLIDTNRSRACHPCSPMKGSKRCWGESSEDDCSLRTFYCAGGACARCKGRPLPTDCCHQ 239
DB 180 ALTLIDTNRSRACHPCSPMKGSKRCWGESSEDDCSLRTFYCAGGACARCKGRPLPTDCCHQ 240
OY 240 CAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNREGRYTFGASCVTAC 299
DB 240 CAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNREGRYTFGASCVTAC 300
OY 300 PYNLTSDVGSCTLVCPHNOEYTAEDGTORCEKSKPCARVCYGLGMELREVRATYSAN 359
DB 300 PYNLTSDVGSCTLVCPHNOEYTAEDGTORCEKSKPCARVCYGLGMELREVRATYSAN 360
OY 360 NIOEFAGCKKIFGSLAFLEPSPFDGPNASNTAPLOPEQLQVFETLEETIGLYISAMPDSL 419
DB 360 NIOEFAGCKKIFGSLAFLEPSPFDGPNASNTAPLOPEQLQVFETLEETIGLYISAMPDSL 420
OY 420 PDLVSFQNLQVIRIGRIHNGAYSILTQGLISWLGRLSRLRELSGGLAIHHNTHLCFVHT 479
DB 420 PDLVSFQNLQVIRIGRIHNGAYSILTQGLISWLGRLSRLRELSGGLAIHHNTHLCFVHT 480
OY 480 VPMDOLFNRPHOALLHTANRPEDCEVGEGLACHQOLCARHGWGPPTQVNCISOFLSGQSC 539
DB 480 VPMDOLFNRPHOALLHTANRPEDCEVGEGLACHQOLCARHGWGPPTQVNCISOFLSGQSC 540
OY 540 VEECRVLOGLPREYVNAHRCPCHEPCOPQNGSVTCFEPADQCVACAHYKDPPECVAC 599
DB 540 VEECRVLOGLPREYVNAHRCPCHEPCOPQNGSVTCFEPADQCVACAHYKDPPECVAC 600
OY 600 CPSSGVPDLSTYMPIMKFPDEEGACOPPCINCHTSCVDLDDKCGPAEQRASPLT----- 652
DB 601 CPSSGVPDLSTYMPIMKFPDEEGACOPPCINCHTSCVDLDDKCGPAEQRASPLT----- 660

```

```
QY 653 ----- 652
Db 661 GVLLELLVVVGIILIKRRROKIRKTYMRLLQETELVEPLTPSGAMPNQAOMRIKETE 720
QY 653 ----- 652
Db 721 LRKVKLGSGAFSTYKGIWIPDGENKIPVAIKVIRENTSPKANKETLDEAYVMAVG 780
QY 653 ----- 652
Db 781 PYVSRLLGICLTSTVOLVQLMPYGLLDHYREHGRGLSGODLLMVCQIAKMSYLEYD 840
QY 653 ----- 652
Db 841 RLVRHDLARNVLYKSPNHVKITDGLARLLDIDETEXHADGKVPKMMALESIILRRF 900
QY 653 ----- 652
Db 901 THOSDWSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGERLPORPCTIDVYMIWKCW 960
QY 653 ----- 652
Db 961 MIDSECRPRRELVSERSMARDPQRFVYIQLNEDLGSPSPMDSTFYRSLLEDMDKGLVD 1020
QY 684 AEETLVPOQGFPCPPAPGAGMVAHHRH 712
Db 1021 AEETLVPOQGFSPDPTPTGTSTAHRRH 1049
```

```
RESULT 11
US-09-870-759-118
; Sequence 118, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERNAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870, 759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: Patent version 3.1
; SEQ ID NO 118
; LENGTH: 1260
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-870-759-118
```

```
Query Match 81.1%; Score 3205; DB 9; Length 1260;
Best Local Similarity 57.9%; Pred. No. 2,36-218;
Matches 608; Conservative 34; Mismatches 70; Indels 338; Gaps 3;
```

```
QY 1 MELALCRWGLLALLPPGASTOYCTGTDMLRLPASPEFHLDMLRHLYGCGOYVGNL 60
Db 4 MELAAWCRWGLLALLPPGASTOYCTGTDMLRLPASPEFHLDMLRHLYGCGOYVGNL 63
QY 61 ELTFLPTNASLFLQDIOEVGYVLIANOVROYLQRLIVRGTOLEFEDNYALAVLDNG 120
Db 64 ELTFLPTNASLFLQDIOEVGYVLIANOVKRYPLQRLIVRGTOLEFEDNYALAVLDNR 123
QY 121 DPLNNTPTVT-GASPGGLRELQSLTELKGLVLIQNPOLCYODTILMKDIFHKNQL 179
Db 124 DPGQNVASFTGRTPEGLRELQSLTELKGLVLIQNPOLCYODTILMKDIFHKNQL 183
QY 180 ALTIDINRSACHPCSPMCKSGRWGESSDQSLRTVCAGGACRGKPLPTDCCHEQ 239
Db 184 APVIDIDINRSACHPCSPMCKSGRWGESSDQSLRTVCAGGACRGKPLPTDCCHEQ 243
QY 240 CAAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDTPESMNPBGRYTFGASCVATC 299
Db 244 CAAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDTPESMNPBGRYTFGASCVATC 303
```

```
QY 300 PYNLTSDVSGCTLVCPRLHNOETAEADGTORCEKSKPCANVCYGLGMEHLREVAVTSA 359
Db 304 PYNLTSTEVGSCCTLVCPRPNNQEVTAEDGTORCEKSKPCANVCYGLGMEHLRGARATISD 363
QY 360 NIOEPAGCKRIEGLAFLPESFDGDPASNTAPLQPEOLQVETLEITGYLIYISAMPDSL 419
Db 364 NVOEFDGCKRIEGLAFLPESFDGDPASNTAPLQPEOLQVETLEITGYLIYISAMPDSL 423
QY 420 PDLVFNQIAYIRGRILHNGAYSLTQIGLISWLGRLSLRELSGLALIHNTILCFVHT 479
Db 424 RDLVFNQIAYIRGRILHNGAYSLTQIGLISWLGRLSLRELSGLALIHNTILCFVHT 483
QY 480 VPMDQLFRNPQALLHTANRDEDE-CVGEGLACIQLACRGKMGKRPQCVNCSOFLRQ 538
Db 484 VPMDQLFRNPQALLHTANRDEDE-CVGEGLACIQLACRGKMGKRPQCVNCSOFLRQ 543
QY 539 ECVEECRYLQGLPREYVYARHCLPCHPCOPONSVCYFGEADQCAACAHYKDRPCVA 598
Db 544 ECVEECRYLQGLPREYVYARHCLPCHPCOPONSVCYFGEADQCAACAHYKDRPCVA 603
QY 599 RCPGVRPDLTYMPIMKFPDEEGACQPCPINCTHSCVDLDKGCAPGASPLT----- 652
Db 604 RCPGVRPDLTYMPIMKFPDEEGACQPCPINCTHSCVDLDKGCAPGASPLT----- 663
QY 653 ----- 652
Db 664 EGVLLFLVVVGIILIKRRROKIRKTYMRLLQETELVEPLTPSGAMPNQAOMRIKET 723
QY 653 ----- 652
Db 724 ELRKVKLGSGAFSTYKGIWIPDGENKIPVAIKVIRENTSPKANKETLDEAYVMAVG 783
QY 653 ----- 652
Db 784 SPYVSRLLGICLTSTVOLVQLMPYGLLDHYREHGRGLSGODLLMVCQIAKMSYLEYD 843
QY 653 ----- 652
Db 844 RLVRHDLARNVLYKSPNHVKITDGLARLLDIDETEXHADGKVPKMMALESIILRR 903
QY 653 ----- 652
Db 904 THOSDWSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGERLPORPCTIDVYMIWKC 963
QY 653 ----- 652
Db 964 WIMSECRPRRELVSERSMARDPQRFVYIQLNEDLGSPSPMDSTFYRSLLEDMDKGLVD 1020
QY 683 DAETLVPOQGFPCPPAPGAGMVAHHRH 712
Db 1024 DAETLVPOQGFSPDPTPTGTSTAHRRH 1053
```

```
RESULT 12
US-09-854-356-14
; Sequence 14, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheysen, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854, 356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 14
```

LENGTH: 1256  
TYPE: prt  
ORGANISM: Mus sp.  
FEATURE:  
OTHER INFORMATION: mouse HER-2/neu protein  
US-09-854-356-14

Query Match 80.7%; Score 3189.5; DB 9; Length 1256;  
Best Local Similarity 57.8%; Pred. No. 2.9e-217;  
Matches 606; Conservative 36; Mismatches 70; Indels 337; Gaps 2;

OY 1 MELALCRMGGLLALLPGAASTGYCTGDMKRLRPASPEHLDMLRHLGYOCQYVQGNL 60  
DB 1 MELAMCMGFLALLSPGAAGTGYCTGDMKRLRPASPEHLDMLRHLGYOCQYVQGNL 60  
OY 61 ELTYLPTNASLSFLDDIOEVGYLIAHNOVRQVPLQRLRIYRGTOLEFEDNTALAVLNDG 120  
DB 61 ELTYLPTNASLSFLDDIOEVGYLIAHNRVKNVPLQRLRIYRGTOLEFEDKVALAVLNDR 120  
OY 121 DPLNN-TTPVVGASPGRLRELOLRSLTELKGVLIQNRNPOLCYODTILMKDIFKKNQL 179  
DB 121 DPLDVTTAAPGRTRPEGRLRELOLRSLTELKGVLIKGNPOLCYODMYLKMVKVLRKNQL 180  
OY 180 ALTLIDTNRSRACHPSPCKSGRCMGSESDQSLTRTVACGAGCARCKGRLPTDCCHEQ 239  
DB 181 APVMDTNRSRACHPSPCAPLTCCKNHCWGSESDQSLTGTICSGCARCKGRLPTDCCHEQ 240  
OY 240 CAAGCTGPRKHSDDLCLHFNHSGICELHCPALVTYNTDTFESMPNDEGRYTFGASCVTAC 299  
DB 241 CAAGCTGPRKHSDDLCLHFNHSGICELHCPALVTYNTDTFESMLNDEGRYTFGASCVTTC 300  
OY 300 PYNLTSTVGSCTLVCPRLHNDVYTAEDGTQRECKSKPCARCYGLGMEHLREVAAYTSA 359  
DB 301 PYNLTSTVGSCTLVCPRLHNDVYTAEDGTQRECKSKPCAGCYGLGMEHLREVAAYTSD 360  
OY 360 NIOEFAGCKKIFGSLAFLPESFDGDPASNTAPLOPEOLOVETLEITGYLITISAMPDLSL 419  
DB 361 NIOEFAGCKKIFGSLAFLPESFDGNPSSGVAPLRKPHLQVETLEITGYLITISAMPESF 420  
OY 420 PDLSTFQNLQVIRGRILHNGAYSLTLOGIGISWLGRLSLRELGSGLALIHNTHLCPVHT 479  
DB 421 PDLSTFQNLQVIRGRILHNGAYSLTLOGIGISWLGRLSLRELGSGLALIHNTHLCPVHT 480  
OY 480 VPMDOLFNPRLHNTANRPEDCEVGEGLACHOLCARGHCMPGPTOCVNCISOFLRGDE 539  
DB 481 VPMDOLFNPRLHNTANRPEDCEVGEGLACHOLCARGHCMPGPTOCVNCISOFLRGDE 540  
OY 540 CVEECRVLOGLPREYVNAHNLCPHRECOPONGSVTCEFPADQCVACAHYKDPFCVYAR 599  
DB 541 CVEECRVWKGLPREYVNRKHLCPHRECOPONGSVTCEFPADQCEACAHYKDSVCYAR 600  
OY 600 CPSSGKRPDLSTWPIKFPDEBAGCOPCPINCTHSCVDLDDKCPAEQORASPLT----- 652  
DB 601 CPSSGKRPDLSTWPIKFPDEEGICOPCPINCTHSCVDLDERGCPAEORASPVTFIATVV 660  
OY 653 ----- 652  
DB 661 GYLLELLIIVYIGLILKRRRQKIRKTYMRLLQETELVEPLTPSGAVPNQAOIRLIKETE 720  
OY 653 ----- 652  
DB 721 LKLLKLVSGAAGTYVKGIMIPDGENVKIPYAIKVLRENTSPKANKELIDEAYVMAAGS 780  
OY 653 ----- 652  
DB 781 PYVSHLIGILSTVQVLTOLMPYGCLDHWREHGRIGLSODLNMVQIAKMSYLEEV 840  
OY 653 ----- 652  
DB 841 RLVDHDLAARNVLYKSPNHVKITDFGLARLLDIDETEVHADGKVP IKMALESILRRRF 900  
OY 653 ----- 652

DB 901 THQSDVWSGYVTWELMTFGAKPYDGIAPAREIPDLLEKGERLPQPICTIDVYIMVMKCM 960  
OY 653 -----SQNEDIGPASPLDSTYRSLLEDDMGDLVD 683  
DB 961 MIDSECRPRFRELWSEFSRMARDPQRFVYIQONEDLGPPSPMDSTYRSLLEDDMGDLVD 1020  
OY 684 AEEYLVPOOGFPCPDPAFGAGMVHRRH 712  
DB 1021 AEEYLVPOOGFSPDPALGTGSTAHRHR 1049

RESULT 13  
US-09-854-356-8

; Sequence 8, Application US/09854356  
; Patent No. US2002017567A1  
; GENERAL INFORMATION:  
; APPLICANT: Cheever, Martin A.  
; APPLICANT: Cheyssen, Dirk  
; APPLICANT: Corixa Corporation  
; APPLICANT: SmithKline Beecham Biologicals S. A.  
; TITLE OF INVENTION: HER-2/neu Fusion Proteins  
; FILE REFERENCE: 014058-009810PC  
; CURRENT APPLICATION NUMBER: US/09/854,356  
; CURRENT FILING DATE: 2001-05-09  
; PRIOR APPLICATION NUMBER: US 09/493,480  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: US 60/117,976  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8

LENGTH: 654  
TYPE: prt  
ORGANISM: Rattus sp.

FEATURE:  
OTHER INFORMATION: extracellular domain (ECD) of rat HER-2/neu  
US-09-854-356-8

Query Match 78.7%; Score 3110.5; DB 9; Length 654;  
Best Local Similarity 85.5%; Pred. No. 4.8e-212;  
Matches 558; Conservative 32; Mismatches 62; Indels 1; Gaps 1;

OY 1 MELALCRMGGLLALLPGAASTGYCTGDMKRLRPASPEHLDMLRHLGYOCQYVQGNL 60  
DB 1 MELAMCMGFLALLSPGAAGTGYCTGDMKRLRPASPEHLDMLRHLGYOCQYVQGNL 60  
OY 61 ELTYLPTNASLSFLDDIOEVGYLIAHNOVRQVPLQRLRIYRGTOLEFEDNTALAVLNDG 120  
DB 61 ELTYLPTNASLSFLDDIOEVGYLIAHNRVKNVPLQRLRIYRGTOLEFEDKVALAVLNDR 120  
OY 121 DPLNN-TTPVVGASPGRLRELOLRSLTELKGVLIQNRNPOLCYODTILMKDIFKKNQL 179  
DB 121 DPLDVTTAAPGRTRPEGRLRELOLRSLTELKGVLIKGNPOLCYODMYLKMVKVLRKNQL 180  
OY 180 ALTLIDTNRSRACHPSPCKSGRCMGSESDQSLTRTVACGAGCARCKGRLPTDCCHEQ 239  
DB 181 APVMDTNRSRACHPSPCAPLTCCKNHCWGSESDQSLTGTICSGCARCKGRLPTDCCHEQ 240  
OY 240 CAAGCTGPRKHSDDLCLHFNHSGICELHCPALVTYNTDTFESMPNDEGRYTFGASCVTAC 299  
DB 241 CAAGCTGPRKHSDDLCLHFNHSGICELHCPALVTYNTDTFESMLNDEGRYTFGASCVTTC 300  
OY 300 PYNLTSTVGSCTLVCPRLHNDVYTAEDGTQRECKSKPCARCYGLGMEHLREVAAYTSA 359  
DB 301 PYNLTSTVGSCTLVCPRLHNDVYTAEDGTQRECKSKPCARCYGLGMEHLREVAAYTSD 360  
OY 360 NIOEFAGCKKIFGSLAFLPESFDGDPASNTAPLOPEOLOVETLEITGYLITISAMPDLSL 419  
DB 361 NIOEFAGCKKIFGSLAFLPESFDGDPSSGIAPLRQLOVETLEITGYLITISAMPDLSL 420  
OY 420 PDLSTFQNLQVIRGRILHNGAYSLTLOGIGISWLGRLSLRELGSGLALIHNTHLCPVHT 479  
DB 421 PDLSTFQNLQVIRGRILHNGAYSLTLOGIGISWLGRLSLRELGSGLALIHNTHLCPVHT 480



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2003, 14:44:35 ; Search time 32.7407 Seconds  
(without alignments)  
4480.838 Million cell updates/sec

Title: US-09-854-356-7  
Perfect score: 3954  
Sequence: 1 METALACRMGLLALLPFGA.....GFCCPDPAAGAGVHHRRH 712

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_21:\*  
1: sp.\_archaea:\*  
2: sp.\_bacteria:\*  
3: sp.\_fungi:\*  
4: sp.\_human:\*  
5: sp.\_invertebrate:\*  
6: sp.\_mammal:\*  
7: sp.\_mhc:\*  
8: sp.\_organelle:\*  
9: sp.\_phage:\*  
10: sp.\_plant:\*  
11: sp.\_rodent:\*  
12: sp.\_virus:\*  
13: sp.\_vertebrate:\*  
14: sp.\_unclassified:\*  
15: sp.\_virus:\*  
16: sp.\_bacteriaph:\*  
17: sp.\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3446.5	87.2	1259	6	018735
2	1874.5	47.4	419	4	090K79
3	1540	38.9	1209	11	090X70
4	1533.5	38.8	655	11	090WPF5
5	1533.5	38.8	1210	11	090EP98
6	1517.5	38.4	643	11	090ERV6
7	1348.5	34.1	1328	13	P79754
8	1346.5	34.1	1165	13	09YH40
9	1156.5	29.2	527	13	09W6F6
10	1155	29.2	527	13	090836
11	1030.5	26.1	1433	5	09B1H9
12	1028.5	26.0	478	11	090SF0
13	971.5	24.6	599	13	090SH2
14	887	22.4	176	11	0923V5
15	734	18.6	331	4	090BD7
16	723	18.3	149	6	090B66

17	656	16.6	1368	5	023821	023821 caenorhabdi
18	621	15.7	1717	5	026566	026566 schistosoma
19	620.5	15.7	1193	5	09Y1X8	09Y1X8 ephyattia f
20	406.5	10.3	1362	13	09PY24	09PY24 xenopus lae
21	404.5	10.2	150	6	09B684	09B684 oryctolagus
22	395	10.0	151	6	09B685	09B685 oryctolagus
23	366.5	9.3	1368	13	08UW85	08UW85 paralicthy
24	366.5	9.3	1472	5	09U5H8	09U5H8 bombyx mori
25	363.5	9.2	366	5	026569	026569 schistosoma
26	362	9.2	1369	13	08UW86	08UW86 paralicthy
27	359	9.1	1358	13	073798	073798 xenopus lae
28	350.5	8.9	136	11	099PK1	099PK1 mus musculus
29	345.5	8.7	1418	13	093457	093457 scophthalmu
30	340.5	8.6	1671	5	09NUV5	09NUV5 biophalar
31	339.5	8.6	410	11	063720	063720 rattus norv
32	338	8.5	412	4	08WYV0	08WYV0 homo sapien
33	331	8.4	334	5	026567	026567 schistosoma
34	331	8.4	342	5	026568	026568 schistosoma
35	328	8.3	149	11	060494	060494 cavia sp. e
36	326.5	8.2	1245	13	09YGH8	09YGH8 scophthalmu
37	325	8.2	1418	13	08UW83	08UW83 paralicthy
38	322	8.1	469	11	063721	063721 rattus norv
39	316	8.0	1371	11	09QVW4	09QVW4 rattus norv
40	315	8.0	1412	13	08UW84	08UW84 paralicthy
41	307	7.8	2144	5	09VD94	09VD94 drosophila
42	302.5	7.7	1749	5	08T0M6	08T0M6 echinococcu
43	296	7.5	946	5	09VU04	09VU04 drosophila
44	284.5	7.2	116	6	09B687	09B687 oryctolagus
45	281	7.1	367	11	08R2X1	08R2X1 mus musculus

## ALIGNMENTS

RESULT 1  
ID 018735 PRELIMINARY; PRT; 1259 AA.  
AC 018735;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Erdb-2.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_Taxid=9615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Yokota H.;  
RT "CDNA cloning of erdb-2 from canine mammary gland."  
RL Submitted (OCT-1997) to the EMBL/GenBank/DBS databases.  
DR EMBL; AB008451; BAA23127.1; -.  
DR HSSP; P11362; IFGK.  
DR InterPro; IPR002048; EF-hand.  
DR InterPro; IPR000494; EGFR\_L\_domain.  
DR InterPro; IPR000719; Euk\_Pkinase.  
DR InterPro; IPR002174; Furin-like.  
DR InterPro; IPR001245; Tyr\_Pkinase.  
DR InterPro; IPR004019; YLP\_motif.  
DR Pfam; PF00757; Furin-like; 1.  
DR Pfam; PF00069; Pkinase; 1.  
DR Pfam; PF01030; Recep\_L\_domain; 2.  
DR Pfam; PF02757; YLP; 2.  
DR ProDom; PD000001; Euk\_Pkinase; 1.  
DR SMART; SM00261; FY; 3.  
DR SMART; SM00219; TYKc; 1.  
DR PROSITE; PS00018; EF\_HAND; UNKNOWN; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
KW ATP-binding; Transferase; Tyrosine-protein kinase.  
SQ SBOUENCE 1259 AA; 137989 MW; E37364D9C4ACD46 CRC64;





DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE Epidermal growth factor receptor.  
 GN EGFR.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NCBI\_TaxID=10116;  
 RX MEDLINE=90258888; PubMed=2342466;  
 RA Petch L.A., Harris J., Raymond V.W., Blasband A.J., Lee D.C.,  
 RA Earp H.S.;  
 RT "A truncated, secreted form of the epidermal growth factor receptor is  
 RT encoded by an alternatively spliced transcript in normal rat tissue.";  
 RL Mol. Cell. Biol. 10:2973-2982(1990).  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FISHER; TISSUE=LIVER;  
 RA Petch L.A.;  
 RL Submitted (NOV-1991) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FISHER; TISSUE=LIVER;  
 RA Cutcliffe K., Dawson T.L., Earp H.S.;  
 RL Submitted (NOV-1991) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; M37394; AAF14008.1; -.  
 DR HSSP; P11362; 1FCR.  
 DR InterPro: IPR000494; EGFR\_L\_domain.  
 DR InterPro: IPR000719; Euk\_kinase.  
 DR InterPro: IPR002174; Furin-like.  
 DR InterPro: IPR001245; Tyr-kinase.  
 DR Pfam; PF00757; Furin-like; 1.  
 DR Pfam; PF00069; kinase; 1.  
 DR Pfam; PF01030; Recep\_L\_domain; 2.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR Prodom; PD000001; Euk-kinase; 1.  
 DR SMART; SM00261; F0; 3.  
 DR SMART; SM00219; Tyrc; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 KW ATP-binding; Receptor; Transferase; Tyrosine-protein kinase.  
 SQ SEQUENCE 1209 AA; 134891 MW; 96FE7F6CC1B7773 CRC64;

Query Match 38.9%; Score 1540; DB 11; Length 1209;  
 Best Local Similarity 43.6%; Pred. No. 6,7e-130;  
 Matches 291; Conservative 101; Mismatches 240; Indels 36; Gaps 9;

QY 3 LAALCWMGLLLALPPGA-ASTOVCTGTDMKRLPASPTHLDMLRHLYOCVOVGNTE 61  
 DB 15 LAALCAAG-----GALEEKVVCOGTSNRLTQLTGFEDHSLQRMFNCEVVLGNLE 66  
 QY 62 LTYLPNASTLSFODIOEYGVYLIHNOVROYQLRLRVGTSTOFEDYVALAVLDND 121  
 DB 67 ITTVQNNYDLSFLKTIQEVAGYVLIHNTVERIPLENLQIRKNALIENTIALAVLSN- 124  
 QY 122 PLNNTPTVYGASPGRLRLQLRSLTEILKGVLIQNRNPOLCYODITLMDIFKNNQAL 181  
 DB 125 -----GYNKGTGRELPNRLQELIIGAVFNSNNPILCNMETIOWNRIV-QDVELSN 175  
 QY 182 TLIDTNRS-RACHPCSPMKSGKNCNCESESDQSILRTYACAGGCA-RCKGRLPTDCCHQ 239  
 DB 176 MSMDVVRHLTGCKPCSPNCSGSCNGEGENCCKLRKIIICAQCSRRCRSPSCDCHQ 235  
 QY 240 CAAGCTGPKHSDCLACILHFNHSGICELHCPALVTYNTDFESMPNPEGRTFGASCVTAC 299  
 DB 236 CAAGCTGPKHSDCLVCHRRDEATCKDCTCPRLPLMYNTTYQMDVNPBGKTSFATCYKAC 295  
 QY 300 PNYTSLTDVSGCTLVCPHNOEVTAEDGTQRCCKSCPCARVCYGIAMEHLREAVRTSA 359

DB 296 PRNYVTVDHSGCVACGPDYEEV-EEEDGVSKCKKCDGCPCKRVCNGIGIEFKDTLSINAT 354  
 QY 360 NIDEPACCKIIPESLAFIPSPFGDPSANTAPLOPQLOVFELEITGTLISAMPDSL 419  
 DB 355 NIHFXYCTAISDGLHLLPFAFGDSFTTRPPLDPRELEIKTKVKEITGTLLOAMPENW 414  
 QY 420 PDSVFNQNVNGLIHLNNGAYSILTQIGISWLSRSLRELSGSLALHNNHLCFVHT 479  
 DB 415 TDLHAFENLEITGRKQKQKQFSLAVVGLNITSLGRSLAKEISDGVYISGNRNLCYANT 474  
 QY 480 VPMDQLFRNPQALHTANRPEDECVGEGLACHQLCARHGCWGPPTOCVNCQSOLRGOE 539  
 DB 475 INMKLFGPNQKTXIMNNNAEKDKATNHCVCPLDSSSECGMPEPTDCSCGNVSRGE 534  
 QY 540 CVEECRYLOGLPREYVNAHCHLCPHPCOPQNSVTCFGEADQVACAHYKDPFCVAR 599  
 DB 535 CVDKCNILREBPREFENSECIOCHPCDLPQTNNICTGCRGNPNCIKCAHYVDGPHCVXT 594  
 QY 600 CPESGVKPDLSYMPKPEPEEGACOPCINCTHSCVDLDKGPAPORASPLTSQEDIG 659  
 DB 595 CPESGIGENNNTL-VMKFADANNVCHLCHANCCTGCGAGPLKGCQPE-----G 641  
 QY 660 PASPLDST 667  
 DB 642 PKIPSIAT 649

RESULT 4  
 ID 09WVF5 PRELIMINARY; PRT; 655 AA.  
 AC 09WVF5;  
 DT 01-NOV-1999 (Tremblrel. 12, Created)  
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE Epidermal growth factor receptor (Epidermal growth factor receptor  
 DE Isoform 3).  
 GN EGFR.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=LIVER;  
 RA Reiter J.L., Threadgill D.W., Danielson A.J., Schell C.,  
 RA Lampard A.L., Balasubramaniam S., Crossley T.O., Magnuson T.R.,  
 RA Mahle N.J.;  
 RT "Alternative transcripts from the Human and Mouse EGFR Genes Encode  
 RT Carboxy-Terminal Truncated Receptors";  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C3H/101, 129/SVJ, AND 129/SVEVYAC;  
 RA Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielson A.J.,  
 RA Sinclair C.S., Pearisall R.S., Green P.J., Yee D., Lampard A.L.,  
 RA Balasubramaniam S., Crossley T.O., Magnuson T.R., James C.D.,  
 RA Mahle N.J.;  
 RT "Comparative genomic sequence analysis and isolation of human and  
 RT mouse alternative Egfr transcripts encoding truncated receptor  
 RT isoforms";  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=LIVER;  
 RX MEDLINE=2108560; PubMed=11217851;  
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Adachi J., Fukuda S.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldi M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Guenichon S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Nomberts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsaki S.,  
 RA Hayashizaki Y.;  
 RA "Functional annotation of a full-length mouse cDNA collection.";  
 RT Nature 409:685-690(2001).  
 RL EMBL; AF124513; AAd4149.1; -;  
 DR EMBL; AF275366; AAG28047.1; -;  
 DR EMBL; AF275364; AAG28047.1; JOINED.  
 DR EMBL; AF275365; AAG28047.1; JOINED.  
 DR EMBL; AK004944; BAB23688.1; -;  
 DR EMBL; AK004883; BAB23641.1; -;  
 DR EMBL; AK004911; BAB23662.1; -;  
 DR MGD; MGI:95294; Egfr.  
 DR InterPro: IPR000494; EGFR\_L\_domain.  
 DR InterPro: IPR002174; Furin-like.  
 DR Pfam; PF00757; Furin-like; 1.  
 DR Pfam; PF01030; Recep\_L\_domain; 2.  
 DR SMART; SM00261; FU; 3.  
 KW Receptor.  
 SO "SEQUENCE 655 AA; 72906 MW; 6B34063B1BC928CB CRC64;

Query Match 38.8%; Score 1533.5; DB 11; Length 655;  
 Best Local Similarity 44.4%; Pred. No. 1.2e-129;  
 Matches 284; Conservative 99; Mismatches 233; Indels 23; Gaps 7;

QY 11 LLLALLPAGA--STVCTGTMKRLPASPETHLDMRLHYQGCGVQVGNLELYPTN 68  
 Db 14 LITLCAAGGALAEKKVCOGTSNRLTQGTEDHFLSLQRMNNECVGLNLEITYVORN 73  
 QY 69 ASLSFLQDIQEVQGYVLAHNOVROPLQRLRIVRGTOLEFEDNVALAVLNDGPNLNTTP 128  
 Db 74 YDLSFLTKTLEQVAGVYLAALNTERIPLENQIINGNALYENTYALALISN----- 124  
 QY 129 VTGASPGGLREQLRLSLREIILKGVLYLQRPOLCYQDTILMKD-----FKKNOIALTLI 184  
 Db 125 -YGTNRRTGLRELPMRNOEILIGAVFSNNPILLCNMDTIOWMDIYQVNFMSNMSDL--- 180  
 QY 185 DTNRSRACHPSCPMCKSGRCWSESDCSLRTVYACAGCA-RCKGPLPTDCCRCQCAAG 243  
 Db 181 -QSHSSCPKCDPSCPNSGCGGEGNOKLTKITCAOQCSHRCKGRSPDCCNQCAAG 239  
 QY 244 CTGPKHSDCLACLHFNHSGICELHCPALVYNTDTFESMPNPEGRYTFGASCVTACPNY 303  
 Db 240 CTGPRESCLVQKQFQDQATCKDTCPLMLXNPYYQMDVNPDEKYSFGATCVKKCPNRY 299  
 QY 304 LSTVSGCTLYCPRLHNOVETADGTQRCCKSPCARVCYGLGMHLEAVANTANIOE 363  
 Db 300 VVTHGSCVRCAGPDYUV-EDDGIRKCKCDGPKRCVONGIGIEFQDTLSINATNIKH 358  
 QY 364 FAGCKKIGSLAFLEPSEFDPASNTAPLQPOLQVFELEIRTEYLYISAMPDPLDS 423  
 Db 359 FKYCAISGDLHILPVAKKGDSFTTPPLDRELEIKTVEITLFTLLOAMPDWTIDH 418  
 QY 424 VFONLOVIRGRILNGAYSLTLOGISLWLGRLSRLRELSGSLALHNHTLCLFVTVPM 483  
 Db 419 AFENLEIRGRTRKQHGQSLAVVGLNITSLGRSLKELSGDVIILSGRNLCYANTIMK 478  
 QY 484 QLFNRPHOALLHTANRPDECVGEGSLACHQLCARHGCHSGPFGTQCVNCSQFLRGCEV 543  
 Db 479 KLFGPNOKTKIMNRAKDKAVNVHVCNPLCSSGCGPREPDCVSCQNSRGRECEK 538  
 QY 544 CRVLQGLREVYVNAHNCBCHPECOPOGNSVTCFGRPADOCVACAHYKDPFCVAKRPSG 603  
 Db 539 CNLIEGREREVENSECIOCHPECLPQAMNITICTGRPNCLQCAHYIDGPHCVKTCPSAG 598  
 QY 604 VKPDLSTYPMIKFPDEBAGACOPCPINCTHSCVLDLDDKGC 642

Db 599 IMGENTL-VWXYADANNVCHLCHANCYGCAGPGLQGC 636

## RESULT 5

QY 09EP98 PRELIMINARY; PRT; 1210 AA.  
 AC 09EP98;  
 DT 01-MAR-2001 (TREMblrel. 16, Created)  
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE Epidermal growth factor receptor isoform 1.  
 GN EGFR.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C3H/101, 129/SVJ, AND 129/SVEVTAJ;  
 RA Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielson A.J.,  
 RA Sinclair C.S., Pearall R.S., Green P.J., Yee D., Lampland A.L.,  
 RA Balasubramaniam S., Crossley T.O., Magnuson T.R., James C.D.,  
 RA Maibale N.J.;  
 RT "Comparative genomic sequence analysis and isolation of human and  
 RT mouse alternative Egfr transcripts encoding truncated receptor  
 RT isoforms.";  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J;  
 RA Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielson A.J.,  
 RA Schehl C., Pearall R.S., Green P.J., Yee D., Lampland A.L.,  
 RA Balasubramaniam S., Crossley T.O., Magnuson T.R., James C.D.,  
 RA Maibale N.J.;  
 RT "Comparative genomic sequence analysis and isolation of human and  
 RT mouse alternative Egfr transcripts encoding truncated receptor  
 RT isoforms.";  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 RT Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF275366; AAG28045.1; -;  
 DR EMBL; AF275364; AAG28045.1; JOINED.  
 DR EMBL; AF275365; AAG28045.1; JOINED.  
 DR EMBL; AF275367; AAG24386.1; -;  
 DR HSSP; P11362; IEGK.  
 DR MGD; MGI:95294; Egfr.  
 DR InterPro: IPR000345; Cytochrome\_bbind.  
 DR InterPro: IPR000494; EGFR\_L\_domain.  
 DR InterPro: IPR000719; Euk\_Pkinase.  
 DR InterPro: IPR002174; Furin-like.  
 DR InterPro: IPR002290; Ser\_thr\_Pkinase.  
 DR InterPro: IPR001245; Tyr\_Pkinase.  
 DR Pfam; PF00757; Furin-like; 1.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR Pfam; PF01030; Recep\_L\_domain; 2.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR PRODOM; PD000001; Euk\_Pkinase; 1.  
 DR SMART; SM00261; FU; 5.  
 DR SMART; SM00220; S\_TKC; 1.  
 DR SMART; SM00219; TYRK; 1.  
 DR PROSITE; PS00190; CYTOCHROME\_C; UNKNOWN.1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS0011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 KW ATP-binding; Receptor; Transferrase.  
 SO SEQUENCE 1210 AA; 134840 MW; 62CD021C9DE32E18 CRC64;

Query Match 38.8%; Score 1533.5; DB 11; Length 1210;  
 Best Local Similarity 44.4%; Pred. No. 2.6e-129;  
 Matches 284; Conservative 99; Mismatches 233; Indels 23; Gaps 7;

QY 11 LLLALLPAGA--STVCTGTMKRLPASPETHLDMRLHYQGCGVQVGNLELYPTN 68  
 Db 14 LITLCAAGGALAEKKVCOGTSNRLTQGTEDHFLSLQRMNNECVGLNLEITYVORN 73



RA Gellner K., Brenner S.;  
 RT "Analysis of 148 kb of genomic DNA around the wnt1 locus of Fugu  
 rubripes.";  
 RL Genome Res. 9:251-258(1999).  
 DR EMBL: AF056116; AAC34391.1; -.  
 DR HSSP: P11362; ITCG.  
 DR InterPro: IPR000494; EGFR\_L\_domain.  
 DR InterPro: IPR000719; Euk\_kinase.  
 DR InterPro: IPR002174; Furin-like.  
 DR InterPro: IPR001245; Tyr\_kinase.  
 DR Pfam: PF00757; Furin-like; 1.  
 DR Pfam: PF00069; pkinase; 1.  
 DR Pfam: PF01030; Recep\_L\_domain; 2.  
 DR ProDom: PD000001; Euk\_pkinase; 1.  
 DR SMART: SM00261; FU; 3.  
 DR SMART: SM00219; TYRK; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR ATP-binding; Transferase.  
 KW SEQUENCE 1328 AA; 148613 MW; A333039258B647E9 CRC64;

Query Match 34.1%; Score 1348.5; DB 13; Length 1328;  
 Best Local Similarity 43.4%; Pred. No. 1.9e-112;  
 Matches 276; Conservative 87; Mismatches 242; Indels 31; Gaps 15;

QY 9 WGLLALPP--GAATG-----VCTGTDMLRLPASPEHIDMLRHLVGGCGOVGNLEL 62  
 Db 4 WRLIMCVASRLRAASSTQTOAVCGPTONGSLSTGSGQENQNLNDRKKEIGETIENLEI 63  
 QY 63 TYLPNTASLFLQDIOEVGYVLIANNOVROVPLORLIRIVGTQLFEDNYALAVLDNDP 122  
 Db 64 TOIESNMDFSLKTRREYLYLIAMNHFOELPLQGLVHNGNSLYERRPLSYVLN--- 120  
 QY 123 LNTPTVGAAPGGLREQLSLFELKGVLIQANPOLCYODTILKMDIRHKNQALTL 182  
 Db 121 ---YPKG--PSGLNQIGLNLTEILDGVQIINNKILRGPVWYWDII--RNNDAEIE 173  
 QY 183 LIDTRSRACHPCSPMCKSGRCWSESSDCSLRTVCAGCG--ARCKGPLPTDCHECA 241  
 Db 174 IQFNGERGVCH--KSC--GNYCWMGKQDQCLITFTVCAPQCNDCPFTSPRDCHECA 229  
 QY 242 AGCGPKASDCLACLHFNHSGICELHCPALVYNTDTFESMPNREGRTFGASCVTACPY 301  
 Db 230 AGCKGPLDTDFACRLFNDSGACVQCQPTLIYNKQTFOMETNPNAKYQSGICVSCPT 289  
 QY 302 NYLSTDVSCSLVCPCLHNOEYTAEDGTOR--CEKSKPCARVCYGLGMHLEVRVTSAN 360  
 Db 290 HFV--VDGSSCVSVCPDKMEV--ERGSORQCELCSGCLPKVCEGTGAE---QROTVDSSN 343  
 QY 361 IOEFGACKRIGSLAFLPESFDPASNTAPLQPEOLQVFTLEBITGYLYISAMPDILP 420  
 Db 344 IDSFNCTKIGQSLHFLVTGLIGDDFKNVPLDAKLEVFVREITDIILNQSPKELN 403  
 QY 421 DLSVFNQIVNRLNGLNGLSLTLOGIGISWLGIRSLRELISGLALIHNTHLCEVHTV 480  
 Db 404 DLSVFSSLTITQGSLSLFRFSIMWNRIPLTLSLGRLSREISDSGVYISQNAHLCHYHTV 463  
 QY 481 PMDLFRPH--QALLHTANRPDECVGEGGLACHOICARGHOCWGPPTOCVNSQGLRQGE 539  
 Db 464 NWTOLFRSRRANSLSNRPMARBCVADGRVCDPLCSGSGWGPEDDCLSCRNSRHGT 523  
 QY 540 CVEBCRVLOGLRREVYNAR--CLPCHPEPCOPONGSVTCFGEADQCVACAHYKDPFPCA 598  
 Db 524 CVACGHNFSIGIRPREFAGLNGVACVACHPECKPOTGKASTGAGADCMACSTFRDPRYCMS 583  
 QY 599 RCPGCVKPDLSYPMIKRPPDEGACQPCPINCTHSC 634  
 Db 584 SCPAGVN--DGEKGLIFKPPNREGHCEPCNONCTGCG 618

RESULT 8  
 Q9YH40 PRELIMINARY; PRT; 1165 AA.

AC Q9YH40;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Receptor tyrosine Kinase proto-oncogene.  
 GN XMRK.  
 OS Xiphophorus xiphidium.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha;  
 OC Cyprinodontiformes; Poeciliidae; Xiphophorus.  
 OX NCBI\_Taxid=8086;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-RIO PURIFICATION;  
 EX MEDLINE-98241172; PubMed-9582016;  
 RA Dunitzljovic N., Winkler C., Wellbrock C., Gomez A., Duschl J.,  
 RT Altschmid J., Scharlt M.;  
 RT "Activation of the Xmrk proto-oncogene of Xiphophorus by  
 overexpression and mutational alterations.";  
 RL Oncogene 16:1681-1690(1998).  
 RP [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-RIO PURIFICATION;  
 RA Scharlt M.;  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U53471; AAD10500.2; -.  
 DR HSSP: P11362; ITCG.  
 DR InterPro: IPR000345; CytC\_heme\_bind.  
 DR InterPro: IPR000494; EGFR\_L\_domain.  
 DR InterPro: IPR000719; Euk\_kinase.  
 DR InterPro: IPR002174; Furin-like.  
 DR InterPro: IPR001899; Gram\_pos\_anchor.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR Pfam: PF00757; Furin-like; 1.  
 DR Pfam: PF00069; pkinase; 1.  
 DR Pfam: PF01030; Recep\_L\_domain; 2.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR ProDom: PD000001; Euk\_pkinase; 1.  
 DR SMART: SM00261; FU; 3.  
 DR SMART: SM00219; TYRK; 1.  
 DR PROSITE: PS00190; CYTOCHROME\_C; UNKNOWN\_1.  
 DR PROSITE: PS00343; GRAM\_POS\_ANCHORING; UNKNOWN\_1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 2.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 KW ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.  
 SQ SEQUENCE 1165 AA; 129614 MW; 7F7EE38D8771A74E CRC64;

Query Match 34.1%; Score 1346.5; DB 13; Length 1165;  
 Best Local Similarity 42.1%; Pred. No. 1.9e-112;  
 Matches 273; Conservative 91; Mismatches 258; Indels 27; Gaps 12;

QY 1 MELALCRWGLLALPPG-AAST-----QVCTGTDMLRLPASPEHIDMLRHLVGGCGOV 55  
 Db 4 LLELEL-----LLLLLSIGRCCSTDPDRKVCQGSNQMTM--LDNHYLKMKKKMYSGCNV 56  
 QY 56 VQGNLELYLPLTNASTSLQDIOEVGYVLIANNOVROVPLQRLIRIVGTQLFEDNYALA 115  
 Db 57 VLENLEITYTQENDSLSTQSIQEVGYVLIAMNEVSTIPLVNRLLRGNLVGKNEFL 116  
 QY 116 VLNDPPLNNTPTVYASPGGLREQLSLFELKGVLIQANPOLCYODTILKMDIFHK 175  
 Db 117 VMSNVQK--NPSSP--DYQVGLKQQLSLNLEILISGKVKYSHNPLLCVETIINMMDIVDK 173  
 QY 176 NNQALTLIDTRSRACHPCSPMCKSGRCWSESSDCSLRTVCAGCG--ARCKGPLPTD 234  
 Db 174 TSNPTMNLIIPAFEROCCDPCGCVNGSCWAPRGHCKFTKLCAEQCNRCRGPRIID 233  
 QY 235 CCHQCAAGCTGPKHSDCLACLHFNHSGICELHCPALVYNTDTFESMPNREGRTFGAS 294  
 Db 234 CCNEHAGCGTGPATDCLACRFDNDGTCKDTPCPKPIYDIVSHOVVADNPINIKITYTGA 293

Oy		295	CVTACPNVYLTSDVSGSTLVCPLHNOEVTEDGTQRECKSKPCARVCYLGMENHLEVR	354
Db		284	CVKRCPNSNYVTE-GAACVRCSSAGMLEVD-ENGKRCKPCDGYCPRKCVDSIGISLSNTI	351
Oy		355	AVTSANTIOEFAGCKKIGCSLAFLPESFDGCPASNTAFLQEPDLQVEPTLEEITGLYISA	414
Db		352	AVNSTINIGSFNCNKINGDIILANNSPEDGRPHKIGMPDEHLMLNLTVAETIEITGLYIMW	411
Oy		415	WPDSLPLDSLFPQNLQVIRGLIHNGAYS-LTLQGLGISWTGLSLRLSRLSGLALIHNNTH	473
Db		412	WPEMNTSLSYQNEIIETRGRTTESRGFSFVVVOYSHLOMWLGSLIKSEVASGNVLTKNPQ	471
Oy		474	LCEFYATPMQDLFNHPHQALLHTNANREDCEVGGLACHOLCARGHOMGGPTQCVCVCSQ	533
Db		472	LRTASTINMRLLPFSEDOSSLTDYDRT-----ENQTCCNNESSDGCGWGSPTYCVSC LH	524
Oy		534	FLARGQECVEBECRVYGLGLPREYVANRHCLPCHPECOPONGSVTCFGPEPADOCVCAHAKDP	593
Db		525	VDRGRCGVASCNLQGBPREAQVQVGCRCVCHQEBELVGTDSLVTCGPBPANCSCAHRRQG	584
Oy		594	PFCVACPSPGVKPDLPSTMPITPKPFDEGACQPCPDINCTHSCVIDIDDKGC	642
Db		585	POCIIPCRPHGMLGDGDFL-IWKYADKKMGKQCPCHQONCTOGCSGPGLSGC	632
<b>RESULT 9</b>				
Oy	Q9W6F6		PRELIMINARY;	PRT; 1137 AA.
ID	Q9W6F6			
AC	Q9W6F6			
DT	01-NOV-1999 (TrEMBLrel. 12, Created)			
DT	01-NOV-1999 (TrEMBLrel. 12, Last sequence update)			
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)			
DE	Receptor tyrosine kinase (Fragment).			
GN	ERBB4.			
OS	Gallus gallus (Chicken).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;			
OC	Gallus.			
OX	NCBI_TaxID=9031;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=HINDBRAIN;			
RX	MEDLINE=99263203; PubMed=10328864;			
RA	Dixon M., Lumsden A.;			
RT	"Distribution of neurotulin-1 (nrg1) and erbB4 transcripts in embryonic chick hindbrain.";			
RL	Mol. Cell. Neurosci. 13:237-258(1999).			
DR	HSP, p11362; IFGK.			
DR	InterPro: IPR000494; EGFR_L_domain.			
DR	InterPro: IPR000719; Euk_pkinase.			
DR	InterPro: IPR002174; Furin-like.			
DR	InterPro: IPR001368; TNFR_c6.			
DR	InterPro: IPR001245; Tyr_Pkinase.			
DR	InterPro: IPR004019; YLP_motif.			
DR	Pfam: PF00757; Furin-like; 1.			
DR	Pfam: PF00069; pkinase; 1.			
DR	Pfam: PF01030; Recep_L_domain; 1.			
DR	Pfam: PF02757; YLP_2.			
DR	PRINTS: PR00109; TYRKINASE.			
DR	ProDom: PD000001; Euk_pkinase; 1.			
DR	SMART: SM00261; FU; 3.			
DR	SMART: SM00219; TYRK; 1.			
DR	PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.			
DR	PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.			
DR	PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.			
DR	PROSITE: PS00652; TNFR_NFR_1; UNKNOWN_1.			
KW	Kinase; Tyrosine-protein kinase.			
FT	NOM_TER	1		
SQ	SEQUENCE	1137 AA; 127927 MW; 40616436587DC084F CMC64;		
<b>Query Match</b>				
		29.28; Score 1156.5; DB 13; Length 1137;		
<b>Best Local Similarity</b>		42.6%; Pred. No. 2.6e-95;		

Matches	207	Conservative	80	Mismatches	186	Indels	7	Gaps	6
QY	161	LCYQDTILMKDIEFKNNQDLATLIDFNRSRACHPCSPMKCSGRGCESSDCQSLTRIVC	220						
Db	3	LCFADTTHMODIVRPMNPMASNFETLPTNGSSGCGRCHKSCGTG - RCWGPTEHNCQTLTKTYC	61						
QY	221	AGGC - ARCKGRLPTCCCHGOCAGCTGPRKHSCLACIHNHSGICELHOPALVTYNTDIF	279						
Db	62	AEQCGRCYCPVPSOCHRECCAGSGSGPKDTCFACAMNNDSGACVYQCPQTFVYNPTTF	121						
QY	280	ESMPNDEGRYTCAGCAVCAFPYNYLSTVDVSGCTLVCPLINDQETAEADTORCKCSKPCA	339						
Db	122	QLEHNHNAKYTGACVCKKCPHNFY - VBSSCVRRACPSKMEV - BENGKMKCKPCTDICP	179						
QY	340	RVCYGLGMEHLREVAVTYSANIQEFAGCKIFGSLAFIPESFDGDPASNTADLPQDQLOV	399						
Db	180	KACDIDIGTGLSVASQTVSSNIDKFTNCKTNGNLIFLTGTGHDGPHYHTIAAINPEKLNLI	239						
QY	400	FETLEITIGLYLISAMPDLPDLSEONLOVYRGRILHGAVALSLTLOGISMLGRSLR	459						
Db	240	FQTVETIEIGYLIQSWPEMDPFRFVSNLVTIGRALSGLSLILIKOOGITSLQPSQSK	299						
QY	460	ELGSGIALIHNTHTLCEFYVTVPMDDLFRRPHQALLHTANRPDECEVGEGLACHQLCARGH	519						
Db	300	QISAGNIITYTDSNLCYHTHTVMTSLTFSPQSKTYIHRKKAKENCTADCMVNECLSSDG	359						
QY	520	CWGPPTQCVNCSQPLRGOECVECEGRVLQGLREXYVMAHCLPCHEQCP - ONGSVTCFG	578						
Db	360	CWGPPTQDCLSCRFIRGTCTIESCNLYDXGEFRFANGSVCECDPQCEKMDNMITYCG	419						
QY	579	PEADQVACAHYKDDPPFCVARKPSSGVKRPDLSTMPYWKPFPEDEGACQPCPINTCHSG	634						
Db	420	PGPDHCTCKCFHKDPCNVCCKPCDGLQAGNSF - IFKYADEDRCHPCHPNCTQGC	473						
RESULT 10									
ID	Q90836	PRELIMINARY;	PRT;	527	AA.				
AC	Q90836;								
DT	01-NOV-1996 (TREMBLrel. 01, Created)								
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)								
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)								
DE	EGF/TGF-alpha receptor precursor.								
GN	C-ERBB.								
OS	Gallus gallus (Chicken).								
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;								
OC	Gallus.								
OX	NCBI_TaxID=9031;								
RN	(1)								
RP	SEQUENCE FROM N.A.								
RX	MEDLINE=92123214; PubMed=1732751;								
RA	Flickinger T.W., Maibhe N.J., Kung H.-J.;								
RT	"An alternatively processed mRNA from the avian c-erbB gene encodes a								
RT	soluble, truncated form of the receptor that can block ligand-								
RT	dependent transformation."								
RL	Mol. Cell. Biol. 12:883-893(1992).								
DR	EMBL: M7637; AAA48759.1; -								
DR	InterPro: IPR000494; EGFRL_domain.								
DR	InterPro: IPR002174; Furin-like.								
DR	Pfam: PF01030; Recep_L-domain; 2.								
DR	SMART: SM00261; FU; 2.								
KW	Receptor; Signal.								
FT	SIGNAL	1	28	POTENTIAL.					
FT	CHAIN	29	527	EGF/TGF-ALPHA RECEPTOR.					
SQ	SEQUENCE	527	AA;	58353	MM;	764564	AABC095298	CRC64;	
Query Match									
		Best local similarity		44.0%		Pred. No. 1.3e-95;			
Matches		227;		Conservative		87;		Mismatches 182;	
						Indels		20;	
								Gaps	
11 LLLALL									

```

Db 20 LLLLLLGRAVLCASVEEKKVCGSTNNKLTQLGHVEDHFTSLQRYNNCEVVLINLETYY 79
Qy 66 PPNASISFLQDIOEYGVYLIANOVROYPLQRLRYRGOLFEDYVYALAVLNGRPLNN 125
Db 80 EHNRLDTLTKTQEVYGVYLIANMVDVPLLENQIIRGVYLIADNSFALAVLSNYH-MNK 138
Qy 126 TPVYVGSFGLRELOLRSTTEILKGVLIQORNPOLCYODTILMKDIFHKNNOLATLID 185
Db 139 TQ-----GLRELPRMKRLSELINGVKISNNPKLCNMDTYLMDIITDSRK-PLTVLD 189
Qy 186 -TNRSAPCHPCSPMKSGKSCWGESSEDCSLRTVCAGCA-RCKGPLPTDCHEQCAAG 243
Db 190 FASNLSSCKCHNCTREDHCWGEQNCOTLTFKVICAOCCSGRCGRKVPSDCCNCCAG 249
Qy 244 CGPKHSDCLACHENHSIGCELCFALTYNNDTFESMNPREGRTFGASCYTACPYNY 303
Db 250 CGPRSDCLACKKEDDATTCKDTCPLVLYNFTYQMDVNPREGKYSFGATCVRECPHN 309
Qy 304 LSTDVGSCTLVCPRLHNOEYTAEDGTORCEKSKPCARVCYGLMEHLREYRAVTSANIOE 363
Db 310 VYDHGSCVRSCTDTYEV-EEGVYRKCKKCDGLSKVCNGLISGLKGLISNATNIDS 368
Qy 364 PACCKIFESLAFPEFSGDPASNTAPLOPELOVFEETLEITGYLYISAMPDSLPLDS 423
Db 369 FNCKTKINDVSLIPYAFGLDAFTKTLPLDPKLDVFRYKELISGLFIQAMPDNATDLY 428
Qy 424 VFONLOVIRGLIHNGAYSLTLQGLISWGLRSLRELGSGLALIHNTNLCFVHPVPM 483
Db 429 AENLEIIRGRTOHGOYSLAVYNLKIQSLGLSLKEISDGLAIKNNKLCYADTMNR 488
Qy 484 QLEFRNPQALLHTANRPEDECVGEGSLACHOLCARGH 519
Db 489 SLFATOSQKTIIQNRRKNDCC-SKSVCPAFAKAH 522

RESULT 11
Q9BIH9 PRELIMINARY: PRT; 1433 AA.
AC Q9BIH9;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, last sequence update)
DE 01-JUN-2002 (Tremblrel. 21, last annotation update)
DE Putative epidermal growth factor receptor (Fragment).
GN EGFR.
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
OC Anopheles.
OX NCBI_taxid=7165;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SU4;
RA Lycett G.J.;
RT "Cloning, expression and localisation of the Anopheles gambiae
RT epidermal growth factor receptor."
RT Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AJ301655; CAC35008.1; -.
DR HSSP; P1362; IFGK.
DR InterPro; IPR000345; Cytc_heme_bind.
DR InterPro; IPR000494; EGFR_L_domain.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002174; Furin-like.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF01030; Recep_L_domain; 2.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00261; FU; 7.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TYRKc; 1.

```

```

DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN.4.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR RECEPTOR.
KW NON_TER
SQ SEQUENCE 1 159585 MW; E3D9D88967724F07 CRC64;

Query Match
Best Local Similarity 33.8%; Pred. No. 8.3e-84; Length 1433;
Matches 233; Conservative 112; Mismatches 258; Indels 87; Gaps 21;

Qy 26 CTGTDMKRLPLASPETHLMRLHYOGCOVVOGNLELTPNASTSFLQDIOEYGVY 85
Db 1 CIGTNGRMVSPANREHYNNKRLNDRYNGCYVGNLEITMIONTNDLNFQIHRVGY 60
Qy 86 IAHNOVROYPLQRLRYRGOLF-----EDNTALAVLNGRPLNNTPTVYVGSFGLREL 140
Db 61 ISLYDLPOVILPRLQIIRGRITFKLNKWEAVGLFY-----SFSHMTL 104
Qy 141 QLRLTEILKGVLIQORNPOLCYODTILMKDI-FHKNNOLATLIDTNRSPACHPCSPMC 199
Db 105 ELPALNDILGSGVGFNNYNLCHKSNWEILLAPQTSKOTTFNFSPEFYVCPCHPSC 164
Qy 200 KGSRCWGESSEDCSLRTVCAGCA-RCKGPLPTDCHEQCAAGCTGPKHSDCLACH 257
Db 165 EVG-CWGECAHNQVRSKLNCSPOSGRCFGKPRECCHLFCAGCGTGTGSDCLACKN 223
Qy 258 FHSIGCELCFALTYNNDTFESMNPREGRTFGASCYTACPYNTLSDVGSCTLVCL 317
Db 224 FYDDGVCKQECPRMQLNPTNFWERNPDGKAYGATVYRKCP-EHLKDNACVYKCPK 282
Qy 318 HNOEYTAEDGTORCEKSKPCARVCYGLMEHLREYRAVTSANIOEFACCKIFGSLAF 377
Db 283 GKMPQNSE-----CVPCKGVCPKTCGEGIVH-----SDNIGYKXCTIIESGLE 329
Qy 378 PESFGDPPASNT-----APLOPELOVFEETLEITGYLYISAMPDSLPLDSVFNQ 429
Db 330 DQSFDEFOOVYTNFSPGPRYIKIDPDLREVFSTYKTEITFINQAHHPFTTLNFRNLE 389
Qy 430 VIRGRILHNGAY-SLTLQGLISWGLRSLRELGSGLALIHNTNLCFVHPVPM 488
Db 390 VVGGRQLKNTLASVYIVTSLKSLKSLKRNSSISYILESDLCFVEDIDWSEIKS 449
Qy 489 PHQALLHTANRPEDECVGEGSLACHOLCARGHCGWGPPTCCVNCQSLRQGEVEEGRVQ 548
Db 450 SDBEVVQNNRNATHEHEGMECSKAGCKGKRECELECKNNKYKCKLDSCK--- 506
Qy 549 GLPREY-VNARHCLPCHRECPQONGSVTFCGPBADQCAAHKDPFVCAKPSGVKPD 607
Db 507 SLPRLYSVDSKTCGDCHQCKD---FCYGPENDCSGCMVVKDGRFCVACCP-TKHA 560
Qy 608 LSYMPTKFRPDEGACQCPINCTHSCVDLDDKGC--PAQRASP--LTSQBDGGRASP 663
Db 561 MN-----GTC---INCHKTCV-----GCKGPRDTIAPDGCISCKAIIAGS--- 597
Qy 664 LDSTFYSLLEDDMDGLDAEYIVPOG 693
Db 598 -DAKIERCLMKDESCPDGY-YSDYIVQEEG 625

RESULT 12
Q9SESO PRELIMINARY: PRT; 478 AA.
AC Q9SESO;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-OCT-2001 (Tremblrel. 18, last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, last annotation update)
DE Epidermal growth factor receptor related protein.
GN ERFP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.

```

OX NCBI\_TaxID-10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-GASTRO-DUODENAL MUCOUS;  
RA Yu Y., Moshier J.A., Majumdar A.P.N.;  
RT "Cloning of a novel Egfr-related peptide: A putative negative  
regulator of Egfr."  
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF187818; AAG17037.2; -  
DR InterPro: IPR000494; EGFR\_L\_domain.  
DR InterPro: IPR002174; Furin-like.  
DR Pfam: PF00757; Furin-like; 1.  
DR Pfam: PF01030; Recep\_L\_domain; 2.  
KW Receptor.  
SQ SEQUENCE 478 AA; 53233 MW; CF873A8376C519E5 CRC64;  
  
Query Match 26.0%; Score 1028.5; DB 11; Length 478;  
Best Local Similarity 42.4%; Pred. No. 3,1e-84;  
Matches 208; Conservative 69; Mismatches 167; Indels 47; Gaps 9;  
  
OY 11 LLLALLPPGAA--STQVCTGDMKRLRLPASPETHLMDLRHLHYGGQVQGNLELTYPN 68  
DB 14 LITLALCAAGALEKKVQGSNRLTQLGTEDHFLSLQRMYNCEVVLNLEITYVGRN 73  
OY 69 ASLSFLQDIQEVGYVLLAHNOVROVPLQRLRIYRGTOLEFEDNYALAVLNDGDLNNTTP 128  
DB 74 YDLSTLKTIQEVAGFLIANTVERIPSEDQIIRGNALYENTYALALSN----- 124  
OY 129 VTGASPGGLRLQLSLTEILKGVLLIQNPOLCYQDTILMKDI---FHKNNQALATLI 184  
DB 125 -YGNRTGLRELPHRNLOELIGAVRFNNPILCNMDTIQWRDIQVNFMSMSMDL--- 180  
OY 185 DTNSRACHPCSPKRCGSGESSEDCQSILTRYCAGGCA-RCKGRLPTDCHBQCAAG 243  
DB 181 -QSHSSCPKDPSPGNSCWGGEENQKLTIKICAAQCSHRCGRSPSDCHQCAAG 239  
OY 244 CTGPKHSCLACLPHNHSIGELHCPALVTYNTDTFESMPNPEGRYTGASCVTACPNY 303  
DB 240 CTGPKSCLVQKQKODATKDCDTPMLYNPTTYQDVNPEGRYTGASCVTACPNY 299  
OY 304 LSTDVSGCTLVCPRLHNOETADGTQRCCKSKPCARVCYGLGMEHLREAVTASNIQ 363  
DB 300 VVTDHSGCVRAAGPRDYEV-EEDGIRKCKKCDGPKRKYCNGSIGEFKDTLSINATNIK 358  
OY 364 FAGCKKIGSLAFIPESPDGPASNTAPLOEQLOVFELEITGYLISAMPDLPDLS 423  
DB 359 FRYCTAISGDHLIPVARKGDSFTPTPLDPRELEIKTEIKTGLISLLIQAMPENWTLH 418  
OY 424 VFQNLQVIRGRILHNGAYSLTLQGLI-----SMLGL-RSLRELGSGLALIHNT 472  
DB 419 AFENLEIRGRKQHGQSLAVVGLNITSLPWQPSLSMQAVTRPLHPLAQ----- 469  
OY 473 HLCFVHTVPM 483  
DB 470 -----NRVSWD 475  
  
RESULT 13  
Q9PSH2 PRELIMINARY; PRT; 599 AA.  
AC Q9PSH2;  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
DE Epidermal growth factor receptor (Fragment).  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID-9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94020816; PubMed=8414496;

RA Callaghan T., Antczak M., Flickinger T., Raines M., Myers M.,  
RA Kung H.J.;  
RT "A complete description of the EGF-receptor exon structure:  
RT implication in oncogene activation and domain evolution."  
RL Oncogene 8:2939-2948(1993).  
DR InterPro: IPR000494; EGFR\_L\_domain.  
DR InterPro: IPR002174; Furin-like.  
DR Pfam: PF00757; Furin-like; 1.  
DR Pfam: PF01030; Recep\_L\_domain; 2.  
DR SMART: SM00261; FU\_3  
SQ SEQUENCE 599 AA; 66363 MW; FEAB46D293D991BD CRC64;  
  
Query Match 24.6%; Score 971.5; DB 13; Length 599;  
Best Local Similarity 30.2%; Pred. No. 5,8e-79;  
Matches 222; Conservative 76; Mismatches 180; Indels 257; Gaps 11;  
  
OY 24 QVCTGDMKRLRLPASPETHLMDLRHLHYGGQVQGNLELTYPNASLSFLQDIQEVGY 83  
DB 5 KYCQGTNNKLTQLQGHVEDHFTSLQRMYNCEVVLNLEITYVEHNRDLTFLKTIQEVAGY 64  
OY 84 VLIANQROYVPLQRLRIYRGTOLEFEDNYALAVLNDGDLNNTTPVTGASPGGLRLQL 143  
DB 65 VLIALNMVDVPLLENLQIRGNVLYDNFSFALAVLSNYH-MNKTQ-----GLRELPMK 115  
OY 144 SLTEILKGVLLIQRNPOLCYQDTILMKDIFHKNNQALATLIDTNSRACHPCSPKCGSR 203  
DB 116 RLSEILNGVAKISNNPKLCNMDTVLMDI-----IDTSK----- 150  
OY 204 CWGESSEDCQSILTRYCAGGCA-RCKGRLPTDCHBQCAAGCTGPKHSCLACLPHNHSIGI 263  
DB 151 -----PL----- 152  
OY 264 CELHCPALVTYNTDTFESMPNPEGRYTGASCVTACPNYLSLTDVSGCTLVCPRLHNOEY 323  
DB 153 -----TVDFASNLSSV 164  
OY 324 AEDGTQRCCKSKPCARVCYGLGMEHLREAVTASNIQEFAGCKKIGSLAFIPESPDG 383  
DB 165 EENGYRKKKCDGLSKCKNGSIGELGILGILSINATNIDSKRCKINGDVSILPVALLG 224  
OY 384 DPASNTAPLOEQLOVFELEITGYLISAMPDLPDLSVFQNLQVIRGRILHNGAYSL 443  
DB 225 DAFTYTLPLDKLIDVFRYKVEISGFLLIQAMPDN-ATLYAFENLEIRGRKQHGQYSL 283  
OY 444 TLQGLISMLGRSLRELGSGLALIHNTNLCEFYVTWMDLFRPHALLHTAR----- 499  
DB 284 AVVNLKIQSLGRSLKEISDSDIALMKNNKLCYADTFNMVRSILFATQSQTKIIONRNKND 343  
OY 500 -----PEDECVGEGE-----ACHQLCARG----- 518  
DB 344 CCRKHPACTEDHCHGAGBQNKQTLTKYICAQGCGRGRKGVPSDCCINQCAAGCTGPRE 403  
OY 519 ----- 518  
DB 404 SDCLACRFRDDATCKDTCPLVLYNPTTYQDVNPEGRYTGASCVTACPNYADRHVC 463  
OY 519 -----HNGRPRQVNCQVFLRGQCEVEGRVLOGLPREYVNAKICLCHPECCQN 571  
DB 464 DPLGSDVGCWGPGRPHCSRFKSKVCVACQCNLLDGEPRFEDSKCLCHSSECLVN 523  
OY 572 G----SVTCFGEADQCVACAHYKDPFCVARGSPGVKFDLSYMPIMKPPDEGACQPCPI 628  
DB 524 STAYVTTSSGGRPHCMCAHIFIDGRHCVKACPRAGVAGENDTL-VKMTADANNAQCQLCRP 582  
OY 629 NCTHSCVDLDDKCP 643  
DB 583 NCTRCKKGPGLGCP 597  
  
RESULT 14  
Q923V5 PRELIMINARY; PRT; 176 AA.  
AC Q923V5;  
RX MEDLINE=94020816; PubMed=8414496;



DT 01-DEC-2001 (TREMblrel. 19, Created)  
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
DE Epidermal growth factor receptor (Fragment).  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MISTAR;  
RA Venzani F.M., Soverchia L., Concetti A.;  
RT "Rat HER-2/neu oncogene sequence: unfaithful nucleotide assignment."  
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF393158; AAK76996.1;  
KW Receptor.  
FT NON\_TER 1 1  
FT NON\_TER 176 176  
SQ SEQUENCE 176 AA; 19491 MW; 80C6594C3EBE32A CRC64;

Query Match 22.4%; Score 887; DB 11; Length 176;  
Best Local Similarity 83.5%; Pred. No. 5,1e-72;  
Matches 147; Conservative 11; Mismatches 18; Indels 0; Gaps 0;

Qy 477 VHTVPMQDLFRNPHQALLHTANRPEDCVGEGLCACHQLCARGHCWGPPTQCVNCSQFLR 536  
Db 1 VHTVPMQDLFRNPHQALLHTANRPEDCVGEGLCACHQLCARGHCWGPPTQCVNCSHFLR 60  
Qy 537 GQECVCECRVLOGLPREYVNAHRLCPHPCOPONGSVTCGPEADOCVACAHYKDDPEFC 596  
Db 61 GQECVCECRVWKGLPREVSVSKRCLPCHPCQPONSSETCFGSEADQCAACAHYKDDSSSC 120  
Qy 597 VARPSSGVKPDLSYMPIMKFPDEEGACQPCPINTHSCVDLDDKCPAEPASPLT 652  
Db 121 VARPSSGVKPDLSYMPIMKFPDEEGICQPCPINTHSCVDLDERGCAEPASPYT 176

RESULT 15  
Q9BUD7  
ID 09BUD7 PRELIMINARY; PRT; 331 AA.  
AC 09BUD7;  
DT 01-JUN-2001 (TREMblrel. 17, Created)  
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)  
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
DE Similar to v-erb-b2 avian erythroblastic leukemia viral oncogene homolog 3.  
OS Homo sapiens (human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PLACENTA;  
RA Strausberg R.;  
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC002706; AAH02706.1;  
DR InterPro; IPR000494; EGFR\_L\_domain.  
DR InterPro; IPR002174; Furin-like.  
DR Pfam; PF00757; Furin-like.1.  
DR Pfam; PF01030; Recep\_L\_domain.1.  
DR SMART; SM00261; FU; 2.  
SQ SEQUENCE 331 AA; 36489 MW; 45B8BEE683FE7E8 CRC64;

Query Match 18.6%; Score 734; DB 4; Length 331;  
Best Local Similarity 44.0%; Pred. No. 7.5e-58;  
Matches 147; Conservative 49; Mismatches 120; Indels 18; Gaps 8;

Qy 10 GLIALLPFGAA--STVCTGTDMKRLRPASPEHMLRLHYOGCOVOCNLELTLYLPT 67  
Db 11 GLIFSLARGSEVGNQAVCEGTGLNGLSVTGDAENQYOTLYKLYRCEVGMNLEIYLTGH 70  
Qy 68 NASLSFLQDIOEVQGYVLIANNOVROVPLQRLRIVRGTOLFEEDVYALAVLDNGDPLNNTT 127

Db 71 NADISFLQWIREVTGYVLVAMNEFSTLPPLNLRVYRGTYQYDGFALFV-----LNYNT 125  
Qy 128 PVTGASPGGLRELOLRSLTEILKGVLIQRPOLCYODITLMKDFEKNQALATLIDTN 187  
Db 126 ----NSSHALRQLRLTQLTETLSGVYIEKNDKICHMDITDMRDIVARD--AEIYKD 178  
Qy 188 RSRACHPCSPCKGSRGWGESSDDQSLFRTVCAGG-ARCKGPLPTDCCHEQCAAGCTG 246  
Db 179 NGRSCPCHVEYCKG-RCWGPSEDCQTLTKTICAPQCNGHCFGPNPQCCHDECAGGCSG 237  
Qy 247 PKHSDCLACLHFNNSGICELHCPALVYVNTDPESMPNPEGRTYFGASCVTACPYNTLS 306  
Db 238 PQDTDFCAFRHFNDSGACVPRCPQPLVYKLTQLEPNPHTKYQYGGVAVASCPHNFV-V 296  
Qy 307 DVGSCITVCPPLHNOEYTAEDGTORCEKCSKPCAR 340  
Db 297 DQTSVACRCPDKKKEVD-KNGLKKECECGGLCPK 329

Search completed: January 13, 2003, 14:48:29  
Job time : 36.7407 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2003, 14:44:10 ; Search time 12.3961 Seconds  
(without alignments)  
3074.904 Million cell updates/sec

Title: US-09-854-356-6  
5078  
Perfect score: 1 MELAALCRKGLLALLPPGA.....TFKGTPTAENPEYGLDVPV 919  
Sequence:

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues  
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4900	96.5	1255	1	ERB2_HUMAN
2	4134	81.4	1257	1	ERB2_RAT
3	4122.5	81.2	1254	1	ERB2_MESAU
4	1643	32.4	1210	1	EGFR_HUMAN
5	1612	31.7	1210	1	EGFR_MOUSE
6	1528.5	30.1	1308	1	ERB4_HUMAN
7	1519	29.9	1308	1	ERB4_RAT
8	1509.5	29.7	703	1	EGFR_CHICK
9	1435	28.3	1342	1	ERB3_HUMAN
10	1374.5	27.1	1339	1	ERB3_RAT
11	1330	26.0	1167	1	XMRK_XIPMA
12	1013.5	20.2	1426	1	EGFR_DROME
13	718.5	14.1	1323	1	LT23_CAEL
14	427	8.4	245	1	ERB2_MOUSE
15	390.5	7.7	1300	1	IRR_MOUSE
16	384	7.6	1363	1	ILPR_BRALA
17	377.5	7.4	1300	1	IRR_CAVPO
18	368	7.2	1382	1	INSR_HUMAN
19	364	7.2	581	1	IRR_RAT
20	362.5	7.1	1383	1	INSR_RAT
21	361	7.1	1372	1	INSR_MOUSE
22	359	7.1	1297	1	IRR_HUMAN
23	353	7.0	1477	1	HHR7_HYDAT
24	326	6.4	1390	1	INSR_AEDAE
25	324	6.4	1607	1	MLPR_LYMET
26	316	6.2	1370	1	IGLR_RAT
27	314	6.2	1373	1	IGLR_MOUSE
28	311.5	6.1	1367	1	IGLR_HUMAN
29	307	6.0	2146	1	INSR_DROME
30	288	5.7	634	1	ERBB_ALV
31	271.5	5.3	1696	1	PKS5_BRACL
32	254.5	5.0	604	1	ERBB_AVIER
33	244.5	4.8	1877	1	PKS5_MOUSE

34	208	4.1	1680	1	FUR2_DROME	P30432 drosophila
35	196	3.9	5179	1	MUC2_HUMAN	Q02817 homo sapien
36	195.5	3.8	937	1	PAC4_RAT	Q63415 rattus norv
37	195	3.8	830	1	SREB_HUMAN	Q14162 homo sapien
38	186.5	3.7	540	1	ERBB_AVIER	P11273 avian eryth
39	184.5	3.7	913	1	PKS5_HUMAN	Q92824 homo sapien
40	180.5	3.6	1877	1	PKS5_RAT	P41413 rattus norv
41	176	3.5	969	1	PAC4_HUMAN	P29122 homo sapien
42	175.5	3.5	2003	1	NTC4_HUMAN	Q99466 homo sapien
43	172.5	3.4	1173	1	TSP1_XENLA	P35448 xenopus lae
44	170	3.3	3312	1	CUR3_HUMAN	Q9nyq7 homo sapien
45	170	3.3	3707	1	PGSM_MOUSE	Q05793 mus musculu

## ALIGNMENTS

RESULT 1  
ERB2\_HUMAN STANDARD: PRT: 1255 AA.  
AC P04626;  
DR 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DI 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Receptor protein-tyrosine kinase erbb-2 precursor (EC 2.7.1.112)  
DE (p185erbB2) (NEU proto-oncogene) (C-erbB-2) (tyrosine kinase-type cell surface receptor HER2) (MLN 19).  
GN ERBB2 OR HER2 OR NGL OR NEU.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
NCBI\_Taxid=9606;  
[1]  
SEQUENCE FROM N.A.  
RX MEDLINE=86118663; PubMed=3003577;  
RA Yamamoto T., Ikawa S., Akiyama T., Semba K., Nomura N., Miyajima N., Saito T., Toyoshima K.;  
RT "Similarity of protein encoded by the human c-erbB-2 gene to epidermal growth factor receptor.";  
RT Nature 319:230-234(1986).  
[2]  
SEQUENCE FROM N.A.  
RX MEDLINE=86070181; PubMed=2999974;  
RA Coussens L., Yang-Feng T.L., Liao Y.C., Chen E., Gray A., McGarath J., Seeburg P.H., Libermann T.A., Schlessinger J., Francke U., Levinson A., Ullrich A.;  
RT "Tyrosine kinase receptor with extensive homology to EGF receptor shares chromosomal location with neu oncogene.";  
RT Science 230:1132-1139(1985).  
[3]  
SEQUENCE OF 737-1031 FROM N.A.  
RX MEDLINE=86016729; PubMed=2995967;  
RA Semba K., Kamata N., Toyoshima K., Yamamoto T.;  
RT "A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-erbB-1/epidermal growth factor-receptor gene and is amplified in a human salivary gland adenocarcinoma.";  
RT Proc. Natl. Acad. Sci. U.S.A. 82:6497-6501(1985).  
[4]  
VARIANTS VAL-654 AND VAL-655.  
RX MEDLINE=93194196; PubMed=8095488;  
RA Ehsani A., Low J., Wallace R.B., Wu A.M.;  
RT "Characterization of a new allele of the human ERBB2 gene by allele-specific competition hybridization.";  
RT Genomics 15:426-429(1993).  
-1- FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX, ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GP30 IS A POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-ALPHA AND AMPHIREGULIN.  
-1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.  
-1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS (POTENTIAL).  
-1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE  
CC RESIDUES (BY SIMILARITY).  
CC -1- POLYMORPHISM: THERE ARE FOUR ALLELES DUE TO THE VARIATIONS IN  
CC POSITIONS 654 AND 655. ALLELE B1 (654-ILE-ILE-655) HAS A FREQUENCY  
CC OF 0.783; ALLELE B2 (654-ILE-VAL-655) HAS A FREQUENCY OF 0.206;  
CC ALLELE B3 (654-VAL-VAL-655) HAS A FREQUENCY OF 0.012.  
CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: M11767; AAA35808.1; -.  
DR EMBL: M11761; AAA35808.1; JOINED.  
DR EMBL: M11762; AAA35808.1; JOINED.  
DR EMBL: M11763; AAA35808.1; JOINED.  
DR EMBL: M11764; AAA35808.1; JOINED.  
DR EMBL: M11765; AAA35808.1; JOINED.  
DR EMBL: M11766; AAA35808.1; JOINED.  
DR EMBL: M11730; AAA35808.1; JOINED.  
DR EMBL: M12036; AAA35978.1; -.  
DR EMBL: X03363; CAA27060.1; -.  
DR PIR: A25491; A25491.  
DR PIR: A24571; A24571.  
DR HSP: P11362; 1FGK.  
DR GeneW: HGNC:3430; ERBB2.  
DR MIM: 164870; -.  
DR InterPro: IPR000494; EGFR\_L\_domain.  
DR InterPro: IPR000719; Euk\_pkinase.  
DR InterPro: IPR001245; Furin-like.  
DR InterPro: IPR004019; Tyr\_pkinase.  
DR Pfam: PF00069; pkinase.1.  
DR Pfam: PF00757; Furin-like.1.  
DR Pfam: PF01030; Recep\_L-domain.2.  
DR Pfam: PF02757; YLP.2.  
DR ProDom: PD000001; Euk\_pkinase.1.  
DR SMART: SM00261; Fv; 3.  
DR SMART: SM00219; TyrKc.1.  
DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP.1.  
DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR.1.  
DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM.1.  
KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;  
KW Transference; Tyrosine-protein kinase; ATP-binding; Phosphorylation;  
KW Polymorphism.  
KW SIGNAL.  
FT CHAIN 1 21 POTENTIAL.  
FT 22 1235 RECEPTOR PROTEIN-TYROSINE KINASE ERBB-2.  
FT DOMAIN 22 652 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 653 675 POTENTIAL.  
FT DOMAIN 676 1255 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 720 987 PROTEIN KINASE.  
FT NP\_BIND 726 734 ATP (BY SIMILARITY).  
FT BINDING 733 733 ATP (BY SIMILARITY).  
FT ACT\_SITE 845 845 BY SIMILARITY.  
FT DISULFID 195 204 BY SIMILARITY.  
FT DISULFID 199 212 BY SIMILARITY.  
FT DISULFID 220 227 BY SIMILARITY.  
FT DISULFID 224 235 BY SIMILARITY.  
FT DISULFID 236 244 BY SIMILARITY.  
FT DISULFID 240 252 BY SIMILARITY.  
FT DISULFID 255 264 BY SIMILARITY.  
FT DISULFID 268 295 BY SIMILARITY.  
FT DISULFID 299 311 BY SIMILARITY.  
FT DISULFID 315 331 BY SIMILARITY.  
FT DISULFID 334 338 BY SIMILARITY.  
FT DISULFID 511 520 BY SIMILARITY.  
FT DISULFID 515 528 BY SIMILARITY.  
FT DISULFID 531 540 BY SIMILARITY.  
FT DISULFID 544 560 BY SIMILARITY.

FT DISULFID 563 576 BY SIMILARITY.  
FT DISULFID 567 584 BY SIMILARITY.  
FT DISULFID 587 596 BY SIMILARITY.  
FT DISULFID 600 623 BY SIMILARITY.  
FT DISULFID 626 634 BY SIMILARITY.  
FT DISULFID 630 642 BY SIMILARITY.  
FT MOD\_RES 1139 1139 BY SIMILARITY.  
FT CARBOHYD 1248 1248 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 124 124 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 187 187 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 530 530 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 571 571 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 629 629 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT VARIANT 654 654 I -> V.  
FT VARIANT 654 654 /FTID-VAR\_004077.  
FT VARIANT 655 655 I -> V.  
FT VARIANT 655 655 /FTID-VAR\_004078.  
FT CONFLICT 1170 1170 P -> A (IN REF. 2).  
FT SEQUENCE 1255 AA; 137909 MW; 39E9DEDA04DCF962 CRC64;  
SQ  
Query Match 96.5%; Score 4900; DB 1; Length 1255;  
Best Local Similarity 73.2%; Pred. No. 1.2e-275; Mismatches 0; Indels 336; Gaps 1;  
Matches 919; Conservative 0; Mismatches 0; Indels 336; Gaps 1;  
QY 1 MELALCRWGLLALLPFGASTGVCTGDMKRLRPSFETHDMLRHLXGCGVVOGNTL 60  
DB 1 MELALCRWGLLALLPFGASTGVCTGDMKRLRPSFETHDMLRHLXGCGVVOGNTL 60  
QY 61 ELTYLPTNASLSFLQDIOEVQGYVLIANOVROYLPDLRLRVRGTOLEEDNYALAVDNG 120  
DB 61 ELTYLPTNASLSFLQDIOEVQGYVLIANOVROYLPDLRLRVRGTOLEEDNYALAVDNG 120  
QY 121 DPLNNTFVPGASPGGLRELOLRSLTEILKGVLIORNPOLCYDDTILMKDIFKNNOLA 180  
DB 121 DPLNNTFVPGASPGGLRELOLRSLTEILKGVLIORNPOLCYDDTILMKDIFKNNOLA 180  
QY 121 DPLNNTFVPGASPGGLRELOLRSLTEILKGVLIORNPOLCYDDTILMKDIFKNNOLA 180  
DB 121 DPLNNTFVPGASPGGLRELOLRSLTEILKGVLIORNPOLCYDDTILMKDIFKNNOLA 180  
QY 181 LTLIDTNRSRACHPCSPWCKSGRSCGSSSEDCQSLRTVCAGGACRCKGRLPTDCHEQC 240  
DB 181 LTLIDTNRSRACHPCSPWCKSGRSCGSSSEDCQSLRTVCAGGACRCKGRLPTDCHEQC 240  
QY 181 LTLIDTNRSRACHPCSPWCKSGRSCGSSSEDCQSLRTVCAGGACRCKGRLPTDCHEQC 240  
DB 181 LTLIDTNRSRACHPCSPWCKSGRSCGSSSEDCQSLRTVCAGGACRCKGRLPTDCHEQC 240  
QY 241 AAGCTGKRHSCLACLFHNSGICELCPALVTYNTDTFESMPNBERYTFGASCVTACP 300  
DB 241 AAGCTGKRHSCLACLFHNSGICELCPALVTYNTDTFESMPNBERYTFGASCVTACP 300  
QY 241 AAGCTGKRHSCLACLFHNSGICELCPALVTYNTDTFESMPNBERYTFGASCVTACP 300  
DB 241 AAGCTGKRHSCLACLFHNSGICELCPALVTYNTDTFESMPNBERYTFGASCVTACP 300  
QY 301 YNYSTVSGSCVLCPLHNOEVTAEQRCCKSCARCYGLGMEHLREVAVTASAN 360  
DB 301 YNYSTVSGSCVLCPLHNOEVTAEQRCCKSCARCYGLGMEHLREVAVTASAN 360  
QY 301 YNYSTVSGSCVLCPLHNOEVTAEQRCCKSCARCYGLGMEHLREVAVTASAN 360  
DB 301 YNYSTVSGSCVLCPLHNOEVTAEQRCCKSCARCYGLGMEHLREVAVTASAN 360  
QY 361 IQEFGAGCKITFGSLAFPESEFGDPAANTAPLQEOLOVETLEITGYLISAMPDPLP 420  
DB 361 IQEFGAGCKITFGSLAFPESEFGDPAANTAPLQEOLOVETLEITGYLISAMPDPLP 420  
QY 361 IQEFGAGCKITFGSLAFPESEFGDPAANTAPLQEOLOVETLEITGYLISAMPDPLP 420  
DB 361 IQEFGAGCKITFGSLAFPESEFGDPAANTAPLQEOLOVETLEITGYLISAMPDPLP 420  
QY 421 DLSYFQNLQVYRGILINNGAYSLTLOGLGISWGLRSLRELGSGLALIHNTHLCPVHTV 480  
DB 421 DLSYFQNLQVYRGILINNGAYSLTLOGLGISWGLRSLRELGSGLALIHNTHLCPVHTV 480  
QY 421 DLSYFQNLQVYRGILINNGAYSLTLOGLGISWGLRSLRELGSGLALIHNTHLCPVHTV 480  
DB 421 DLSYFQNLQVYRGILINNGAYSLTLOGLGISWGLRSLRELGSGLALIHNTHLCPVHTV 480  
QY 481 PMDOLFNRPHOALHTANRPEDCEVGEGLACHQICANGHCMPPTOCVACSKQSLRQEC 540  
DB 481 PMDOLFNRPHOALHTANRPEDCEVGEGLACHQICANGHCMPPTOCVACSKQSLRQEC 540  
QY 481 PMDOLFNRPHOALHTANRPEDCEVGEGLACHQICANGHCMPPTOCVACSKQSLRQEC 540  
DB 481 PMDOLFNRPHOALHTANRPEDCEVGEGLACHQICANGHCMPPTOCVACSKQSLRQEC 540  
QY 541 VEECRVLQGLPREYVNAHCLPCHECOPONGSVTCGPEADQCVACAHYKDPFCVARC 600  
DB 541 VEECRVLQGLPREYVNAHCLPCHECOPONGSVTCGPEADQCVACAHYKDPFCVARC 600  
QY 541 VEECRVLQGLPREYVNAHCLPCHECOPONGSVTCGPEADQCVACAHYKDPFCVARC 600  
DB 541 VEECRVLQGLPREYVNAHCLPCHECOPONGSVTCGPEADQCVACAHYKDPFCVARC 600  
QY 601 PSQVKKPDLSTYPIKPFDEBGAQPCPINCTHSCVDLDDKGCAPAEQASPLTS----- 653  
DB 601 PSQVKKPDLSTYPIKPFDEBGAQPCPINCTHSCVDLDDKGCAPAEQASPLTS----- 653  
QY 601 PSQVKKPDLSTYPIKPFDEBGAQPCPINCTHSCVDLDDKGCAPAEQASPLTS----- 653  
DB 601 PSQVKKPDLSTYPIKPFDEBGAQPCPINCTHSCVDLDDKGCAPAEQASPLTS----- 653  
QY 654 ----- 653  
DB 661 ILVVVLLGVVFGILIKRROOKIRKTYMRLLOETELVEPLTPSGAMPNOAMRIKETEL 720

QY	654	-----	655
Db	721	RKKVKVLGSGAFGIYYIGIMLPIDGENVKIPVAIKLRENTSPKANKEIIDEAIVAGVSP	780
QY	654	-----	653
Db	781	YVSRLLGICLTSTVQLVLTQMLPQGCILDHVRNRRGLSQDILLNMCQILANGMSYLEDVR	840
QY	654	-----	653
Db	841	LVHRDLAARNLVKSPNHVKITDFGLARLDIDETEHADGKVPKIMMALESILRRFT	900
QY	654	-----	653
Db	901	QSDVMSYGVTVWELMTFGAKPYDGIPIAREIDPLEKGERLPPICITIDVIMYKCM	960
QY	654	-----	654
Db	961	IDSECRREFELVSEFSRRAADQRRVYIQNEDLGPASPLDSTFYRSLLEDMDMDLVDA	1020
QY	685	EEYLVPOQGFECDDPAPGAGGVNHRHRSSTRSGGDLTLGLESEEDAPRSLAPSEG	744
Db	1021	EEYLVPOQGFECDDPAPGAGGVNHRHRSSTRSGGDLTLGLESEEDAPRSLAPSEG	1080
QY	745	AGSDVPEGDLGMCMAAGLQSLPHDPSPLQRYSEDPVYLPSETDGYAAPLTCSPQRYV	804
Db	1081	AGSDVPEGDLGMCMAAGLQSLPHDPSPLQRYSEDPVYLPSETDGYAAPLTCSPQRYV	1140
QY	805	NOEDVPRQPPSPRECPPLPAPRAGATLERPKTLSPGKNGVWVDVAFGAVENPEYLPQ	864
Db	1141	NOEDVPRQPPSPRECPPLPAPRAGATLERPKTLSPGKNGVWVDVAFGAVENPEYLPQ	1200
QY	865	GGAAPQHPHPAPSPAFNDLYWDODPERKAPSPSTFKCTPFAENPEYLGIDVPY	919
Db	1201	GGAAPQHPHPAPSPAFNDLYWDODPERKAPSPSTFKCTPFAENPEYLGIDVPY	1255
RESULT 2			
ERB2_RAT			
ID	ERB2_RAT	STANDARD:	PRT: 1257 AA.
AC	P06494;		
DT	01-JAN-1988 (Rel. 06, Created)		
DT	15-DEC-1998 (Rel. 37, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)		
DE	(p185erbB2) (NNU proto-oncogene) (C-erbB-2) (Epidermal growth factor receptor-related protein).		
DE	ERBB2 OR NEU.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
OX	NCBI_TaxID=10116;		
OX	[1]		
RN	SEQUENCE FROM N.A.		
RP	TISSUE=Neuroblastoma;		
RC	TISSUE=Sciatic nerve;		
RC	TISSUE=Sciatic nerve;		
RX	MEDLINE=91222560; PubMed=2025425;		
RA	Lai C., Lemke G.;		
RT	"An extended family of protein-tyrosine kinase genes differentially		
RT	expressed in the vertebrate nervous system.";		
RL	Neuron 6:691-704(1991).		
RN	[3]		
RP	STRUCTURE BY NMR OF 650-668		
RP	MEDLINE=92155181; PubMed=1346763;		
RA	Gullick W.J., Bottomley A.C., Loftis F.J., Doak D.G., Mulvey D.,		

[illegible]

```

FT DISULFID 569 586 BY SIMILARITY.
FT DISULFID 589 598 BY SIMILARITY.
FT DISULFID 602 625 BY SIMILARITY.
FT DISULFID 628 636 BY SIMILARITY.
FT DISULFID 632 644 BY SIMILARITY.
FT MOD.RES 1141 1141 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD.RES 1250 1250 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 532 532 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 573 573 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 631 631 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 661 661 V -> E (IN ONCOGENIC NBD).
SQ SEQUENCE 1257 AA; 138831 MW; 6129264583011402 CRC64;

Query Match 81.4%; Score 4134; DB 1; Length 1257;
Best Local Similarity 62.1%; Pred. No. 2e-231;
Matches 781; Conservative 42; Mismatches 96; Indels 338; Gaps 3;

QY 1 MEALALCRMGILLALLPFGAASVQCTGDMKRLPASPETHLDMRLHLYOGCVOVGNL 60
DB 1 MEALAMCRMGILLALLPFGIAGVQCTGDMKRLPASPETHLDMRLHLYOGCVOVGNL 60
QY 61 ELTYLPNASLFLQDIQVQGYVLAHNOYRYPLORLKRVSTQLEPDNYALAVLDNG 120
DB 61 ELTYVPANASLFLQDIQVQGYVLAHNOYRYPLORLKRVSTQLEPDNYALAVLDNR 120
QY 121 DPLNNTTPVT-GASPGGLRELOLRSLTEILKGVLIQIRNPOLCQDPTILMKDIFHKNNOL 179
DB 121 DPLDNVASTPGRTPRELRELOLRSLTEILKGVLIQIRNPOLCQDPTILMKDIFHKNNOL 180
QY 180 ALTLIDNRSRACHPCSPMKSGSRMGESSEDCSLTRVYACAGCARCKGPLPTDCHEQ 239
DB 181 APVDIDNRSRACHPCSPMKSGSRMGESSEDCSLTRVYACAGCARCKGPLPTDCHEQ 240
QY 240 CAAGCTGPKHSDDLACHLPHNSHGICELHCPALVTYNTDFESMNPGRRTYFGASCYTAC 299
DB 241 CAAGCTGPKHSDDLACHLPHNSHGICELHCPALVTYNTDFESMNPGRRTYFGASCYTAC 300
QY 300 PYNVLSTDVSGCTVCPPLHNOEYVAEDGTORCEKSPKPCAVCGGLMEHLREVRATSA 359
DB 301 PYNVLSTDVSGCTVCPPLHNOEYVAEDGTORCEKSPKPCAVCGGLMEHLREVRATSD 360
QY 360 NIOEFACCKIFGSLAFIPESFDGDPASNTAPLOPEQLOVEFTELEITGYLYISAMPDSL 419
DB 361 NIOEFACCKIFGSLAFIPESFDGDPASNTAPLOPEQLOVEFTELEITGYLYISAMPDSL 420
QY 420 PDLSVFONLOYIRGRIILHNGAYSLTLOGLGISWGLSLRSLGSLALIHNTILCFVHT 479
DB 421 PDLSVFONLOYIRGRIILHNGAYSLTLOGLGISWGLSLRSLGSLALIHNTILCFVHT 480
QY 480 VPMQDLFRNPHQALLHNTANPREDE-CVGEGLACHQLCARGCWPGPTQCNCOSQPLRGQ 538
DB 481 VPMQDLFRNPHQALLHNTANPREDE-CVGEGLACHQLCARGCWPGPTQCNCOSQPLRGQ 540
QY 539 ECVEECVHVLQGLPREYVNAHCHLPCHEPCOPONGSVTCFGEADQCAACAYKPPPCVA 598
DB 541 ECVEECVHVLQGLPREYVNAHCHLPCHEPCOPONGSVTCFGEADQCAACAYKPPPCVA 600
QY 599 RCPGSKPDLSTYMPIMKFPDEGACQPCPINCSTHSCVDLDDKGPAPGRASPLT- 652
DB 601 RCPGSKPDLSTYMPIMKFPDEGACQPCPINCSTHSCVDLDDKGPAPGRASPLT- 660
QY 653 ----- 652
DB 661 VGVLLFLILVVVVGILKRRRQKIRKYMRLLOETELVEPLTSPGAMPOAQRILKET 720
QY 653 ----- 652
DB 721 ELRKVKVLSGAGFTGVKGIWIPDGENVKIPVAIKVLENTSPKANKEILDEAVYVAGVG 780
QY 653 ----- 652

```

```

DB 781 SPYVSRLLGICLITSTVOLVTLQMPYGCILDVHEHNRGLGSODLLMWCVOIAKGMSTLED 840
QY 653 ----- 652
DB 841 VRLVHRDLAARNVLYKSPNNVKTITDEGLANLLIDETEFYHADGKVPITKMALESLIRRR 900
QY 653 ----- 652
DB 901 FTHQSDVMSYGVYVWELMTFGAKPYDGPAREIPDLLEKEGERTLPQPICTIDVYIMVYKC 960
QY 653 ----- 652
DB 961 WMIDSECRPPRRELVSFSSRMADPQRFVYIQNDELPSSPMSTFYSRLLEDMDGDLV 1020
QY 683 DAEEYLVPOQGFCCPDAPAGCMVHHRHSSSTRSGGDLTGLFESEEARSPPLAPS 742
DB 1021 DAEEYLVPOQGFSPDPPTGTSTAHNRHSSSTRSGGDLTGLFESEEARSPPLAPS 1080
QY 743 EGAGSDVFDGDLGGAAGLQSLFTHDPSPLOYSSEDPVPLPSENDGYVAPLTCSPQPE 802
DB 1081 EGAGSDVFDGDLGGAAGLQSLFTHDPSPLOYSSEDPVPLPSENDGYVAPLTCSPQPE 1140
QY 803 YVNPDPVQPPSPRREGPLAARPAAGATLERPTLSGKGVYKDVAFEGAVENPEYLT 862
DB 1141 YVNPSEVQPPSPRREGPLAARPAAGATLERPTLSGKGVYKDVAFEGAVENPEYLT 1200
QY 863 POGGAAPQPPPPAFSPAFDNLVYWDQDPBERGAPSTFGPAENPEYILGIDVPV 919
DB 1201 PRECTASPPHSPAFSPAFDNLVYWDQNSGQPPSPFNFGTPTAENPEYILGIDVPV 1257

RESULT 3
ERB2_MESAU
ID ERB2_MESAU STANDARD; PRT; 1254 AA.
AC 060553;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Receptor protein-tyrosine kinase erbB-2 precursor (BC 2.7.1.112)
DE (p185erbB2) (NEU proto-oncogene) (C-erbB-2).
GN ERB2 OR NEU.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxId=10036;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Nerve;
RX MEDLINE=94193007; PubMed=7908275;
RA Nakamura T., Ushijima T., Ishikawa T.,
RA Yamazaki Y., Ishikawa T.;
RT Cloning and activation of the Syrian hamster neu proto-oncogene.
RT Gene 140:251-255(1994).
CC -|- FUNCTION: ESSENTIAL COMPONENT OF A NEREGULIN-RECEPTOR COMPLEX,
CC ALTHOUGH NEUROGLINS DO NOT INTERACT WITH IT ALONE. GP30 IS A
CC POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-
CC ALPHA AND AMPHIREGULIN (BY SIMILARITY).
CC -|- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
CC tyrosine phosphate.
CC -|- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS
CC (POTENTIAL).
CC -|- SUBCELLULAR LOCATION: Type I membrane protein.
CC -|- PRIM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE
CC RESIDUES.
CC -|- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; D16295; BAA03801.1; -.  
DR HSSP; P1162; IFGK.  
DR InterPro: IPR000494; EGFR\_L\_domain.  
DR InterPro: IPR000719; Euk\_pkinase.  
DR InterPro: IPR002174; Furin-like.  
DR InterPro: IPR001245; Tyr\_pkinase.  
DR InterPro: IPR004019; YLP\_motif.  
DR Pfam: PF00757; Furin-like; 1.  
DR Pfam: PF01030; Recept\_L\_domain; 2.  
DR Pfam: PF02757; YLP; 2.  
DR ProDom: PD000001; Euk\_pkinase; 1.  
DR SMART; SM00261; FO; 3.  
DR SMART; SM00219; TYRKC; 1.  
DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;  
KW Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;  
KW Proto-oncogene; Disease mutation.  
FT SIGNAL 1 21 POTENTIAL.  
FT CHAIN 22 1254 RECEPTOR PROTEIN-TYROSINE KINASE ERBB-2.  
FT DOMAIN 22 652 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 653 675 POTENTIAL.  
FT DOMAIN 676 1254 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 158 368 CYS-RICH.  
FT DOMAIN 472 644 CYS-RICH.  
FT DOMAIN 720 987 PROTEIN KINASE.  
FT NP\_BIND 726 734 ATP (BY SIMILARITY).  
FT BINDING 753 753 ATP (BY SIMILARITY).  
FT ACT\_SITE 845 845 BY SIMILARITY.  
FT DISULFID 195 204 BY SIMILARITY.  
FT DISULFID 199 212 BY SIMILARITY.  
FT DISULFID 236 244 BY SIMILARITY.  
FT DISULFID 240 252 BY SIMILARITY.  
FT DISULFID 255 264 BY SIMILARITY.  
FT DISULFID 268 295 BY SIMILARITY.  
FT DISULFID 299 311 BY SIMILARITY.  
FT DISULFID 315 331 BY SIMILARITY.  
FT DISULFID 334 338 BY SIMILARITY.  
FT DISULFID 511 520 BY SIMILARITY.  
FT DISULFID 515 528 BY SIMILARITY.  
FT DISULFID 531 540 BY SIMILARITY.  
FT DISULFID 544 560 BY SIMILARITY.  
FT DISULFID 563 576 BY SIMILARITY.  
FT DISULFID 567 584 BY SIMILARITY.  
FT DISULFID 587 596 BY SIMILARITY.  
FT DISULFID 600 623 BY SIMILARITY.  
FT DISULFID 626 634 BY SIMILARITY.  
FT DISULFID 642 642 BY SIMILARITY.  
FT MOD\_RES 1139 1139 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
FT MOD\_RES 1247 1247 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
FT CARBOHYD 68 68 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 125 125 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 187 187 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 259 259 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 530 530 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 571 571 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 629 629 N-LINKED (GLCNAC...) (POTENTIAL).  
FT VARIANT 658 658 V -> E (IN ONCOGENIC NEU).  
FT VARIANT 659 659 V -> E (IN ONCOGENIC NEU).  
SQ SEQUENCE 1254 AA; 138252 MW; 974C3791C21F2BE1 CRC64;

Query Match 81.2%; Score 4122.5; DB 1; Length 1254;  
Best local Similarity 61.8%; Pred. No. 9, 2e-231;  
Matches 776; Conservative 49; Mismatches 93; Indels 337; Gaps 2;

QY 1 MELALALCRWGLLALLPPGAAS7OVC7GTDMLRLPASPTHLDMRLHLYQGCGVVOGNTL 60  
DB 1 MELALAMCGMLLALLSPGAS7OVC7GTDMLRLPASPTHLDIYVHLYQCGCVVOGNTL 60

QY 61 ELRYLPNASTLSPLODIOEVGYVYLIAHNOVQVPLQRLRTVRSGLFEDNVALAVLDNG 120  
DB 61 ELRYLPANATLSPLODIOEVGYVYLIAHNOVQVPLQRLRTVRSGLFEDNVALAVLDNR 120  
QY 121 DPLNNTPTVVGASPGSLREVLQSLSTELKGVYLQGNPOLCYODTITMIDIFKNNOLA 180  
DB 121 DPLDNVTATGRDEGLREVLQSLSTELKGVYLIRGNPOLCYODTITMIDIFKNNOLA 180  
QY 181 LTLIDNRSRACPCSPMCKSGRCSGESSSDCSLRTYVAGGACARCKPLPTDCCHQC 240  
DB 181 PVDIDNRSRACPCSPMCKSGRCSGESSSDCSLRTYVAGGACARCKPLPTDCCHQC 240  
QY 241 AAGCTGKHSDCLACLFHNSGICELCPALVYNTFTFESMPREBRTFFGASCTACR 300  
DB 241 AAGCTGKHSDCLACLFHNSGICELCPALVYNTFTFESMPREBRTFFGASCTATCP 300  
QY 301 YNYLSTVSGCTVCPPLHNOEVTAEEDTQCEKSKRCARVCYGLGMEHLREYRAVTSAN 360  
DB 301 YNYLSTVSGCTVCPPLHNOEVTAEEDTQCEKSKRCARVCYGLGMEHLRGARATTSAN 360  
QY 361 IOEPAGCKITFGSLAFLPESFDGDPASNTAPLOPEOLQVEETLEITGYLYISAMPDLP 420  
DB 361 IOEPAGCKITFGSLAFLPESFDGDPASNTAPLOPEOLQVEETLEITGYLYISAMPDSLH 420  
QY 421 DLSYFQNLQYIRGRILHNGAYSLTGLGISTWGLSRLSGSLALIHNTILCYHTY 480  
DB 421 DLSYFQNLQYIRGRILHNGAYSLTGLGISTWGLSRLSGSLALIHNTILCEYHTY 480  
QY 481 PMDOLFNPQALHTANPREDECVSGELACHOLCAHGMGPGPTCVCSOLFNGQEC 540  
DB 481 PMDOLFNPQALHTANPREDECVSGELACHOLCAHGMGPGPTCVCSOLFNGQEC 540  
QY 541 VEECRVLQGLPREVYVNAHCLPCHPECPONGSVTCGPEADOCVACAHKDPFCVARC 600  
DB 541 VEECRVLQGLPREVYVNAHCLPCHPECPONGSVTCGPEADOCVACAHKDPFCVARC 600  
QY 601 PSQVKPPLSTMPITMKPFDEGACQPCPINCTHSCVLDLDDKGCRAEONASLTLS 653  
DB 601 PSQVKPPLSTMPITMKPFDEGACQPCPINCTHSCVLDLDDKGCRAEONASLTLS 653  
QY 654 ----- 653  
DB 654 ----- 653  
QY 661 ILFLVYGVVYGLIKRRRQKIRKYMRLLOETELVEPLTPSGAMPNQAOMILKETEL 720  
DB 661 ILFLVYGVVYGLIKRRRQKIRKYMRLLOETELVEPLTPSGAMPNQAOMILKETEL 720  
QY 654 ----- 653  
DB 654 ----- 653  
QY 721 RKVKVLGSAFGTYVKGKIWIPIGGENYKIPVAIKVIRENTSPKANKETLDEAYVMAIGLSP 780  
DB 721 RKVKVLGSAFGTYVKGKIWIPIGGENYKIPVAIKVIRENTSPKANKETLDEAYVMAIGLSP 780  
QY 654 ----- 653  
DB 654 ----- 653  
QY 781 YVSRLLGICLSTVOLVQOLMPYGCLLDHVREHGRGLSQDLLMNCVOIAKMSYLEDVR 840  
DB 781 YVSRLLGICLSTVOLVQOLMPYGCLLDHVREHGRGLSQDLLMNCVOIAKMSYLEDVR 840  
QY 654 ----- 653  
DB 654 ----- 653  
QY 841 LVHRDLAARVNLKSPNHVKITDFGLARLLDIDETEVHADGKGVPIKWLLESILRRRT 900  
DB 841 LVHRDLAARVNLKSPNHVKITDFGLARLLDIDETEVHADGKGVPIKWLLESILRRRT 900  
QY 654 ----- 653  
DB 654 ----- 653  
QY 901 HQSDVMSYGVVWELMTFGAKPYDGIPIAREIPDLLEKGERLPORICTIDVYIMVCKWM 960  
DB 901 HQSDVMSYGVVWELMTFGAKPYDGIPIAREIPDLLEKGERLPORICTIDVYIMVCKWM 960  
QY 654 ----- 653  
DB 654 ----- 653  
QY 961 IDSECRPRELYSESRKARDPQREYVIONEDGSPSSPLDSTFTYSLDEDDMGDLVDA 1020  
DB 961 IDSECRPRELYSESRKARDPQREYVIONEDGSPSSPLDSTFTYSLDEDDMGDLVDA 1020  
QY 685 EEYLVPQGFCDPAPAGAGVNHRRSSSTSGGDLTLGLPSESEAPRSLAPSEG 744  
DB 685 EEYLVPQGFCDPAPAGAGVNHRRSSSTSGGDLTLGLPSESEAPRSLAPSEG 744  
QY 1021 EEYLVPQGFCDPAPAGAGVNHRRSSSTSGGDLTLGLPSESEAPRSLAPSEG 1080  
DB 1021 EEYLVPQGFCDPAPAGAGVNHRRSSSTSGGDLTLGLPSESEAPRSLAPSEG 1080  
QY 745 AGSDVFDGDLGKAAGLQSLPHHDSPLQRYSEDPTVPLPSETDGVADLTCSPQPEYV 804  
DB 745 AGSDVFDGDLGKAAGLQSLPHHDSPLQRYSEDPTVPLPSETDGVADLTCSPQPEYV 804  
QY 1081 AGSDVFDGDLGKAAGLQSLPHHDSPLQRYSEDPTVPLPSETDGVADLTCSPQPEYV 1140  
DB 1081 AGSDVFDGDLGKAAGLQSLPHHDSPLQRYSEDPTVPLPSETDGVADLTCSPQPEYV 1140

OY 805 NOPDVRRPQPSREGLPAPAPAGATLERKTLSPGKNGVYKDVFAFGAVENPEYLTPQ 864  
 DB 1141 NOPEVRPQPLRPEGLPVPVAPGATLERKTLSPGKNGVYKDVFFGAVENPEYLTPR 1200  
 OY 865 GGAAPOPHPPEAFDFLXYWDDPPERGAPEPTFFKPTPAENPEYGLVPPV 919  
 DB 1201 GGSASQPH-PALCALFAFDNLYWDDPPSRGSPPTTFECTPTAENPEYGLVPPV 1254

RESULT 4  
 EGFR\_HUMAN STANDARD: PRT: 1210 AA.  
 ID EGFR\_HUMAN  
 AC P005133; P06268; Q14225; Q9UMD7; Q9UMD8; Q9UMG5; Q9Z795; C00732;  
 AC 000688; Q9B2S2; Q9H2C9; Q9GZX1; Q9H3C9;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Epidermal growth factor receptor precursor (EC 2.7.1.112) (Receptor  
 protein-tyrosine kinase ErbB-1).  
 GN EGFR OR ERBB1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RX MEDLINE=84219729; PubMed=6328312;  
 RA Ullrich A., Coussens L., Hayflick J.S., Dull T.J., Gray A., Tam A.W.,  
 RA Lee J., Varden Y., Libermann T.A., Schlessinger J., Downward J.,  
 RA Mayes E.L.V., Whittle N., Waterfield M.D., Seeburg P.H.;  
 RT "Human epidermal growth factor receptor cDNA sequence and aberrant  
 RT expression of the amplified gene in A431 epidermoid carcinoma cells.";  
 RT Nature 309:418-425(1984).  
 RL [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RX TISSUE=Placenta;  
 RX MEDLINE=95382857; PubMed=7654368;  
 RA Ilekis J.V., Stark B.C., Scoccia B.;  
 RT "Possible role of variant RNA transcripts in the regulation of  
 RT epidermal growth factor receptor expression in human placenta.";  
 RT Mol. Reprod. Dev. 41:149-156(1995).  
 RL [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RX TISSUE=Placenta;  
 RX MEDLINE=97078686; PubMed=8918811;  
 RA Reltter J.L., Maible N.J.;  
 RT "A 1.8 kb alternative transcript from the human epidermal growth  
 RT factor receptor gene encodes a truncated form of the receptor.";  
 RT Nucleic Acids Res. 24:4050-4056(1996).  
 RL [4]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RX TISSUE=Placenta;  
 RX MEDLINE=97256547; PubMed=9103388;  
 RA Ilekis J.V., Garfili J., Niederberger C., Scoccia B.;  
 RT "Expression of a truncated epidermal growth factor receptor-like  
 RT protein (TEGFR) in ovarian cancer.";  
 RT Gynecol. Oncol. 65:36-41(1997).  
 RL [5]  
 RP SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).  
 RX TISSUE=Placenta;  
 RX MEDLINE=21100872; PubMed=11161793;  
 RA Reltter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielsen A.J.,  
 RA Schell Stinchlir C., Pearsall R.S., Green P.J., Yee D., Lampland A.L.,  
 RA Balasubramaniam S., Crossley T.D., Magnuson T.R., James C.D.,  
 RA Maible N.J.;  
 RT "Comparative genomic sequence analysis and isolation of human and  
 RT mouse alternative EGFR transcripts encoding truncated receptor  
 RT isoforms.";  
 RT Genomics 71:1-20(2001).  
 RL [6]  
 RP SEQUENCE OF 575-687 FROM N.A.  
 RA Reltter J.L., Threadgill D.W., Danielsen A.J., Schell C.M.,  
 RA Lampland A.L., Balasubramaniam S., Crossley T.O., Magnuson T.R.,

RA Maible N.J.;  
 RT "Human and mouse alternative EGFR transcripts encoding only the  
 RT extracellular domain of the receptor.";  
 RT Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.  
 RL [7]  
 RP SEQUENCE OF 713-924 FROM N.A.  
 RX MEDLINE=84196372; PubMed=6326261;  
 RA Lin C.R., Chen W.S., Kruliger W., Stolarsky L.S., Weber W.,  
 RA Evans R.M., Verma I.M., Gill G.N., Rosenfeld M.G.;  
 RT "Expression cloning of human EGF receptor complementary DNA: gene  
 RT amplification and three related messenger RNA products in A431  
 RT cells.";  
 RT Science 224:843-848(1984).  
 RL [8]  
 RP SEQUENCE OF 150-962 FROM N.A.  
 RX MEDLINE=84245835; PubMed=6330563;  
 RA Xu Y.H., Ishii S., Clark A.J.L., Sullivan M., Wilson R.K., Ma D.P.,  
 RA Roe B.A., Merlino G.T., Pastan I.;  
 RT "Human epidermal growth factor receptor cDNA is homologous to a  
 RT variety of RNAs overproduced in A431 carcinoma cells.";  
 RT Nature 309:806-810(1984).  
 RL [9]  
 RP SEQUENCE OF 1028-1210 FROM N.A.  
 RX MEDLINE=85046483; PubMed=6093780;  
 RA Simmen F.A., Gope M.L., Schultz T.Z., Wright D.A., Carpenter G.,  
 RA O'Malley B.W.;  
 RT "Isolation of an evolutionarily conserved epidermal growth factor  
 RT receptor cDNA from human A431 carcinoma cells.";  
 RT Biochem. Biophys. Res. Commun. 124:125-132(1984).  
 RL [10]  
 RP SEQUENCE OF 1-29 FROM N.A.  
 RX MEDLINE=88217333; PubMed=3329716;  
 RA Haley J.D., Whittle N., Bennett P., Kinchington D., Ullrich A.,  
 RA Waterfield M.D.;  
 RT "The human EGF receptor gene: structure of the 110 kb locus and  
 RT identification of sequences regulating its transcription.";  
 RT Oncogene Res. 1:375-396(1987).  
 RL [11]  
 RP SEQUENCE OF 1-29 FROM N.A.  
 RX MEDLINE=91107677; PubMed=1988448;  
 RA Haley J.D., Waterfield M.D.;  
 RT "Contributory effects of de novo transcription and premature  
 RT transcript termination in the regulation of human epidermal growth  
 RT factor receptor proto-oncogene RNA synthesis.";  
 RT J. Biol. Chem. 266:1746-1753(1991).  
 RL [12]  
 RP SEQUENCE OF 1-29 FROM N.A.  
 RX MEDLINE=85270438; PubMed=2991899;  
 RA Ishii S., Xu Y.H., Stratton R.H., Roe B.A., Merlino G.T., Pastan I.;  
 RT "Characterization and sequence of the promoter region of the human  
 RT epidermal growth factor receptor gene.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 82:4920-4924(1985).  
 RL [13]  
 RP SEQUENCE OF 540.  
 RA Kohda D.;  
 RT Submitted (SEP-1997) to the SWISS-PROT data bank.  
 RL [14]  
 RP RECEPTOR ACTIVITY.  
 RX MEDLINE=84191554; PubMed=6325948;  
 RA Mroczkowski B., Mosig G., Cohen S.;  
 RT "ATP-stimulated interaction between epidermal growth factor receptor  
 RT and supercoiled DNA.";  
 RT Nature 309:270-273(1984).  
 RL [15]  
 RP PHOSPHORYLATION.  
 RX MEDLINE=89278137; PubMed=2543678;  
 RA Margolis B.L., Lax I., Kris R., Dombalagian M., Honegger A.M.,  
 RA Hawk R., Givol D., Ullrich A., Schlessinger J.;  
 RT "All autophosphorylation sites of epidermal growth factor (EGF)  
 RT receptor and HER2/neu are located in their carboxyl-terminal tails.  
 RT Identification of a novel site in EGF receptor.";  
 RT J. Biol. Chem. 264:10667-10671(1989).  
 RL [16]





```

QY 811 PDPSPREGLPAPAPAGATL-----ERPKTLSPGKNGVKDVFAGGAVENPEYL-TPQ 864
DB 1096 -----SVP-KRPAGSVQNPVYHNQPLNAPSROPNYOD--PHSAGNPPEYLNTVQ 1143
QY 865 GGAAPQHPAPAPADNLYMDO-----DP-----PERGAPSTFGTPTPAE 908
DB 1144 -----PFCVNSTPDSAPHAQKSHQISLDNPDYQODPFPEAKNGCIFKGS-TAE 1193
QY 909 NPEYL 913
DB 1194 NAEYL 1198

RESULT 5
EGFR_MOUSE
ID EGFR_MOUSE STANDARD; PRT; 1210 AA.
AC 001279;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Epidermal growth factor receptor precursor (EC 2.7.1.112).
GN EGFR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=Liver;
RX MEDLINE=93026370; PubMed=1408137;
RA Ayali A., Skorecki K., Yaron A., Glivol D.;
RT "Promoter region of the murine fibroblast growth factor receptor 2
RT (bek/KGFR) gene.";
RL Oncogene 7:1957-1962(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C, and CD-1; TISSUE=Liver, and Decidua;
RX MEDLINE=93126380; PubMed=7678348;
RA Paria B.C., Das S.K., Andrews G.K., Dey S.K.;
RT "Expression of the epidermal growth factor receptor gene is regulated
RT in mouse blastocysts during delayed implantation.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:55-59(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=Liver;
RA Hibbs M.L.;
RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=B6/C3; TISSUE=Liver;
RX MEDLINE=94170986; PubMed=8125255;
RA Luettike N.C., Phillips H.K., Qiu T.H., Copeland N.G., Earp H.S.,
RA Jenkins N.A., Lee D.C.;
RT "The mouse waved-2 phenotype results from a point mutation in the EGF
RT receptor tyrosine kinase.";
RL Genes Dev. 8:399-413(1994).
RN [5]
RP SEQUENCE OF 1-714 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=91232866; PubMed=2030916;
RA Ayali A., Lax I., Ullrich A., Schlessinger J., Glivol D., Morse B.;
RT "Comparison of EGF receptor sequences as a guide to study the ligand
RT binding site.";
RL Oncogene 6:673-676(1991).
RN [6]
RP SEQUENCE OF 969-1117 FROM N.A.
RC STRAIN=C3H;
RA Eisinger D.P., Serrero G.;
RL Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THE EGF RECEPTOR MEDIATES THE BIOLOGICAL SIGNAL OF EGF,
CC AND ALSO OF TGF-ALPHA, AMPHIREGULIN, HEPARIN-BINDING EGF, GP30 AND
CC VACCINIA VIRUS GROWTH FACTOR (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein

```

```

CC Tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SUBCELLULARS: Binding of EGF to the receptor leads to
CC induction, internalization of the EGF-receptor complex,
CC induction of the tyrosine kinase activity, stimulation of cell DNA
CC synthesis, and cell proliferation.
CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X78987; CAAS5587.1; -
DR EMBL; U03425; AAA17899.1; -
DR EMBL; X59698; CAA42219.1; -
DR EMBL; L06864; AAA53029.1; -
DR HSP; L12608; CAA78249.1; -
DR HSP; P11362; IFCG
DR MGD; MGI:95294; Egfr.
DR InterPro; IPR000494; EGFR_L_domain.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002174; Furin-like.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF01030; Recep_L_domain; 2.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00261; FU; 3.
DR SMART; SM00219; TYRKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR Transmembrane; Glycoprotein; Receptor; Signal; Transferase;
KW Tyrosine-protein kinase; ATP-binding; Phosphorylation; Repeat.
FT SIGNAL 1 24
FT CHAIN 25 1210
FT DOMAIN 25 647
FT TRANSMEM 648 670
FT DOMAIN 671 1210
FT REPEAT 75 300
FT REPEAT 390 600
FT DOMAIN 1028 1071
FT DOMAIN 714 981
FT NP_BIND 720 728
FT BINDING 747 747
FT ACT_SITE 839 839
FT DISULFID 190 199
FT DISULFID 194 207
FT DISULFID 215 223
FT DISULFID 219 231
FT DISULFID 232 240
FT DISULFID 236 248
FT DISULFID 251 260
FT DISULFID 264 291
FT DISULFID 295 307
FT DISULFID 311 326
FT DISULFID 329 333
FT DISULFID 506 515
FT DISULFID 510 523
FT DISULFID 526 535
FT DISULFID 526 555
FT DISULFID 539 555
FT DISULFID 558 571
FT DISULFID 562 579
FT DISULFID 582 591
FT DISULFID 595 617
FT DISULFID 620 628
FT DISULFID 624 636
FT MOD_RES 680 680
FT MOD_RES 1092 1092

```



FT	MOD_RES	1110	1110	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT	MOD_RES	1172	1172	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT	MOD_RES	1197	1197	PHOSPHORYLATION (AUTO-, MAJOR SITE)
FT	CARBOHYD	128	128	(BY SIMILARITY).
FT	CARBOHYD	175	175	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	196	196	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	352	352	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	413	413	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	444	444	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	528	528	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	568	568	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	603	603	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	623	623	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT	19	19	C -> S (IN REF. 2).
FT	CONFLICT	539	539	C -> W (IN REF. 5).
FT	CONFLICT	991	991	L -> F (IN REF. 4).
FT	CONFLICT	1116	1117	HP -> DR (IN REF. 6).
SO	SEQUENCE	1210	AA: 134853 MW: 690E20D46DFD2D2F5 CRC64;	
Query Match				
Best Local Similarity 31.7%; Score 1612; DB 1; Length 1210;				
Matches 378; Conservative 131; Mismatches 313; Indels 454; Gaps 25;				
QY	11	L L L A L P P G A A - - S T O V C T G T D M K L R P A S P E T H L D M R H L Y O G C O V V Q G N L E L Y L P T N	68	
DB	14	L L T L A C A G A L E E K K V O G G T S N R L T Q C T F E D H F L S L Q R M Y N N E E V L L G N L E I Y V O R N	73	
QY	69	A S L S L D I O E V O G V L I A H N O V R P L Q R L R Y R G T O L F E D N A L A V L D N G D P L N T T P	128	
DB	74	Y D L S L K T I O E A G V V L A L T A L T V E R I P L E N L O I R G N A L Y E N T Y A L I L S N - - - - -	124	
QY	129	V T G A S P G L R E L O L R S L E I I K G G V L I O R N P O L C Y O D F I L M K D I - - - - F H K N N O L A L T L I	184	
DB	125	- Y C N R T G L R E L P M N L O E I L I G A V R S N P I L C M D T I O M D I O V N F M S M D L - - - -	180	
QY	185	D T N S R A C H P C S P M C K S R C W G S E S D C O S L T R Y C A G C A - R C K G P L P T D C C H E Q C A A G	243	
DB	181	- Q S H P S C P K D P S C P N S C M G G E E N Q K L I I C A Q C C H R C G R S P S D C H N Q C A A G	239	
QY	244	C T G P R H S P C L A C L H F N H S G I C E L H C P A L V Y N T D F E S M P N E G R Y T F G A S C V T A C P Y N	303	
DB	240	C T G P R E S D C L V C O K Q D E A T C K D T C P R L M L Y N P T Y Q M D V N E G K I S F G A T C V K C P R N Y	299	
QY	304	L S T D V G S C T L V C P L N O E V T A D E Q O R E K S C P A R C Y C L G M H L E V A N Y S A N T O E	363	
DB	300	V Y T D H G S C V R A C G P D Y E V - E E D G I R K K C K O C G P R K Y C N G I G E E F K T L S I N A T N I K H	358	
QY	364	F A G C K I G S L A F L P E S F D P A S T A P L O P E O L O V F T L E I T G Y L I S A M P D S L P D L S	423	
DB	359	F K Y C A I S G D L H I L P V A K G D S F T R P L D R E L I K T V E I T G F L L O A M P M W T D L H	418	
QY	424	V P O N L O V I R G I L N G A V S T L O G I S W I G L R S L R E L G S I A L I H N T H L C F H T V P W D	483	
DB	419	A F E N L E I I R G T K O G O F S L A V G I N T S L G R S L K E I S D G V I I S G R N C A Y A N T I M K	478	
QY	484	O L F R N P H A L L T A N R P E D E V G E G L A C H O L C A R G H C M G P R T O C V N S O L R Q E C E E	543	
DB	479	K L F G P N O K T I M N N R A K E K D A V N H V C N P L S C S G C M G P R D V S C O N Y S R G E C E K	538	
QY	544	C R V L O G L P R E V N A R H C L P R E C O P O N G S Y T C F G P E A D O C V A C A H Y K D P F C V A R C P S G	603	
DB	539	C N I L E G E R E V E N S E C L Q C H P E C L P O A M N I T T G R G P D N C I O C A H I D G P H C V T C P A G	598	
QY	604	V K P D L S Y P I W K P D E G A C O P R I N C H S C V D L D K G C P A - - - - -	644	
DB	599	I M G E N N T L - V M K Y A D A N N V C H L C H A N C Y G C A G P L O G C E V W P S G P K I P S I A T G I V G L L	657	
QY	645	- - - - - E D R A S P L T S Q N E - - - - -	656	
DB	658	F I V V A L G I G L F M R R H I V R R T L R L L Q E R E L V E P L P S G A P N O A H L R I L K E T E F K K I	717	
QY	657	- - - - -	656	

DB	718	K V L G S G A G Y V Y K G L M I P E G E K V I P A I K E L R E A T S P K A N K E I D E A Y V A S Y D N P H V C	777
QY	657	- - - - -	656
DB	778	R L L G I C L S T V O L I T Q L M P R Y G C L D Y V R E N K D I G S Q Y L N M W C V O I A K M N Y L E D R L V H	837
QY	657	- - - - - D L G - - - - -	659
DB	838	R D L A A R N V L V T P O H V K I T D F G L A K L L G A E K E Y H A E G K Y P I K M A L E S I L H R Y T H Q S	897
QY	660	- - - - - P A - - - - -	661
DB	898	D W S Y G V T W E L M F G S K P Y D G I P A S D I S L E K G E R L P O P P I C T I D Y Y M I M V K O M I D A	957
QY	662	- - - - - S P I D S T F F Y R S L E D D M G D L V D A E	686
DB	958	D S R P K F R E L I E F S K M A R D P O R Y L V I O G D E R H N L P S P T D S N F Y R A L M D E E D M E D V D A D E	1017
QY	687	Y L V P O G F C P D P A P A G A G M V H H R R S S T R S G G D L T L G L E P S E E A P R S P L A P S E G A G	746
DB	1018	Y L I P O G F F - - - - - N S P S T - - - - - S R P L L S S L A T	1043
QY	747	S D V E D G L G M G A K G L S L P T H D P S P L O R Y S E D P T V P L P S E T - - D G Y A P L T C S P O P E Y V	804
DB	1044	S N - - - - N S T V A C I N R N G S C R V K E D A F L Q R Y S D P T G A T E D N I D A F L - - - - - P V P E Y V	1093
QY	805	N O P D Y R O P P R P R E G P L P A A R A G A T L - - - - - E R K T L S P K N G V Y K V F A F G A V E N P E	859
DB	1094	N O - - - - - S V P - K R P A G S V O N P V Y H N O P L H A P A R D L I Y O N - - P H S N A V N P E	1137
QY	860	Y L - P P O G G A A P R P P R A F S P A F D L Y W D O - - - - - P R - - - - - P E R G A P P S F K	902
DB	1138	Y L N T A Q - - - - - P T L S S G F N S P A L M I O K G S H O M L D N P D Y O Q D F F P K E T R N G I F K	1188
QY	903	G T P T A E N P E Y L G L D V P	918
DB	1189	G - P T A E N A E Y L R V A P P	1203
RESULT 6			
ERB4_HUMAN			
ID	ERB4_HUMAN	STANDARD;	PRT; 1308 AA.
AC	Q15303		
DT	15-DEC-1998 (Rel. 37, Created)		
DT	15-DEC-1998 (Rel. 37, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Receptor protein-tyrosine kinase erbB-4 precursor (EC 2.7.1.112)		
DE	(p180erbB4) (Tyrosine kinase-type cell surface receptor HER4).		
GN	ERB4 OR HER4.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N. A. (ISOFORM JM-A).		
RC	TISSUE=Breast carcinoma;		
RX	MEDLINE=93189574; PubMed=8383326;		
RA	Piomman G.D., Culouscou J.-M., Whitney G.S., Green J.M., Carlton G.W.,		
RA	Foy L., Neubauer M.G., Shoyab M.;		
RT	"Ligand-specific activation of HER4/p180erbB4, a fourth member of the		
RT	epidermal growth factor receptor family."		
RL	Proc. Natl. Acad. Sci. U.S.A. 90:1746-1750(1993).		
RN	[2]		
RP	SEQUENCE FROM N. A. (ISOFORMS JM-A AND JM-B).		
RC	TISSUE=Fetal brain;		
RX	MEDLINE=97476287; PubMed=9334263;		
RA	Elenius K., Cortas G., Paul S., Choi C.J., Rio C., Piomman G.D.,		
RA	Klagsdorn M.;		
RT	"A novel juxtamembrane domain isoform of HER4/erbB4. Isoform-specific		
RT	tissue distribution and differential processing in response to		
RT	phorbol ester."		
RL	J. Biol. Chem. 272:26761-26768(1997).		

CC -1- FUNCTION: SPECIFICALLY BINDS AND IS ACTIVATED BY NEUREGULINS, NRG-  
CC 2, NRG-3, HEPARIN-BINDING EGF-LIKE GROWTH FACTOR, BETACELLULIN AND  
CC NTAR. INTERACTION WITH THESE FACTORS INDICES CELL DIFFERENCIATION.  
CC NOT ACTIVATED BY EGF, TGF-A, AND AMPHIREGULIN.  
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein  
CC tyrosine phosphate.  
CC -1- SUBUNIT: HOMODIMER OR HETERODIMER WITH EACH OF THE OTHER ERBB  
CC RECEPTORS (POTENTIAL).  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: JM-A (SHOWN HERE) AND JM-B;  
CC ARE PRODUCED BY ALTERNATIVE SPLICING. THE TWO ISOFORMS DIFFER  
CC FUNCTIONALLY IN THEIR RESPONSE TO PHORBOL ESTER: JM-A IS PROCESSED  
CC BUT NOT JM-B. SO, THEY RESPECTIVELY REPRESENT CLEAVABLE AND  
CC NONCLEAVABLE FORMS OF THE RECEPTOR. BOTH ISOFORMS ARE EXPRESSED IN  
CC CEREBELLUM, BUT ONLY THE JM-B ISOFORM IS EXPRESSED IN THE HEART.  
CC -1- TISSUE SPECIFICITY: EXPRESSED AT HIGHEST LEVELS IN BRAIN, HEART,  
CC KIDNEY. IN ADDITION TO SKELETAL MUSCLE, PARATHYROID, CEREBELLUM,  
CC PITUITARY, SPLEEN, TESTIS AND BREAST. LOWER LEVELS IN THYMUS,  
CC LUNG, SALIVARY GLAND, AND PANCREAS.  
CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE  
CC RESIDUES.  
CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: L07868; AAB59446.1; -  
DR HSSP: P11362; IEGR.  
DR GeneW: HGNC:3432; ERBB4.  
DR MIM: 600543; -  
DR InterPro: IPR000494; EGFR\_L.domain.  
DR InterPro: IPR000719; Euk\_pkinase.  
DR InterPro: IPR002174; Furlin-like.  
DR InterPro: IPR001245; Tyr\_pkinase.  
DR InterPro: IPR004019; YLP\_mocif.  
DR Pfam: PF00069; pkinase; 1.  
DR Pfam: PF00757; Furlin-like; 1.  
DR Pfam: PF01030; Recep\_L-domain; 2.  
DR Pfam: PF02757; YLP; 2.  
DR ProDom: PD000001; Euk\_pkinase; 1.  
DR SMART: SM00261; FU; 4.  
DR SMART: SM00219; TYRKC; 1.  
DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;  
KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;  
KW Alternative splicing.  
FT CHAIN 1 25 POTENTIAL.  
FT CHAIN 26 1308 RECEPTOR PROTEIN-TYROSINE KINASE ERBB-4.  
FT DOMAIN 26 651 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 652 675 POTENTIAL.  
FT DOMAIN 676 1308 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 186 334 CYS-RICH.  
FT DOMAIN 496 633 CYS-RICH.  
FT DOMAIN 718 985 PROTEIN KINASE.  
FT NP\_BIND 724 732 ATP (BY SIMILARITY).  
FT BINDING 751 751 ATP (BY SIMILARITY).  
FT ACT\_SITE 843 843 BY SIMILARITY.  
FT DISULFID 189 197 BY SIMILARITY.  
FT DISULFID 193 205 BY SIMILARITY.  
FT DISULFID 213 221 BY SIMILARITY.  
FT DISULFID 217 229 BY SIMILARITY.  
FT DISULFID 230 238 BY SIMILARITY.  
FT DISULFID 234 246 BY SIMILARITY.  
FT DISULFID 249 258 BY SIMILARITY.  
FT DISULFID 262 289 BY SIMILARITY.  
FT DISULFID 293 304 BY SIMILARITY.

FT DISULFID 308 323 BY SIMILARITY.  
FT DISULFID 326 330 BY SIMILARITY.  
FT DISULFID 503 512 BY SIMILARITY.  
FT DISULFID 507 520 BY SIMILARITY.  
FT DISULFID 523 532 BY SIMILARITY.  
FT DISULFID 536 552 BY SIMILARITY.  
FT DISULFID 555 569 BY SIMILARITY.  
FT DISULFID 559 577 BY SIMILARITY.  
FT DISULFID 580 589 BY SIMILARITY.  
FT DISULFID 593 614 BY SIMILARITY.  
FT DISULFID 617 625 BY SIMILARITY.  
FT DISULFID 621 633 BY SIMILARITY.  
FT MOD\_RES 1162 1162 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
FT MOD\_RES 1188 1188 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
FT MOD\_RES 1258 1258 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
FT MOD\_RES 1284 1284 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
FT CARBOHYD 138 138 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 174 174 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 181 181 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 253 253 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 358 358 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 410 410 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 473 473 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 495 495 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 548 548 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 576 576 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 620 620 N-LINKED (GLCNAC...) (POTENTIAL).  
FT VARSPLIC 626 648 NPTSHDCIYVWMTGSHLPQHA -> ICSSTIEDCIGLMD  
FT (IN ISOFORM JM-B).  
SQ SEQUENCE 1308 AA: 146807 MW: 564AB80985DB8761 CRC64;  
Query Match 30.1%; Score 1528.5; DB 1; Length 1308;  
Best Local Similarity 27.7%; Pred. No. 5,5e-81;  
Matches 372; Conservative 132; Mismatches 338; Indels 499; Gaps 25;  
QY 9 WGLLLALPPGAA-----STGYCTGTMKRLRPSPEFHLDMLRHLYQGCGVYQGNLELY 64  
DB 8 WWWWSLVLAAGTVQPSDSQSCAGTEKKLSLSLDEQQYALKRKYENEVEVNGNLEITS 67  
QY 65 LPTNASTSFLQDIOEVQVYLIANOVQVPLRIIVRGTOLEFEDNIALAVLADNDPNTN 124  
DB 68 IEHNRLSLFASRYREVTGYVLVLANOPRYLPLENLAIIKGTKEDEYRYALATLNRKQ 127  
QY 125 NTPVYTGASPGGLREQLSLTEILKGVLIQNPOLCYQDTITLMDKDFHKNQALTLI 184  
DB 128 NF-----GLOELGKLKLTLEILNGGVYDQKFLCYADITFIMODIVRNPWPSMLTY 178  
QY 185 DTRSRACHPGCMKGRSGSSSDQSLPTFVQAGGC-ARCKRPLDPCCHBCAAG 243  
DB 179 STGSSGCGCRHKSCGTG-KCWGPTENHCQTLRTVCAEQDGRKYGVYSDCCHRCAAG 237  
QY 244 CTGPKHSDCIACLPFHNSGICELHCPALVYVNTDFESMNPREGRTFGASCATAPYVY 303  
DB 238 CSFPTKTDGACAGNPFDSGACVYQCQPTVYNTFTQLLEINFAKTTYGAFCKKPRHNF 297  
QY 304 LSTDVGSCITLVCPLHNOEYTAEDGTGRCEKSCPKARVCYGLMEHLREYRAVTSANIOE 363  
DB 298 V-VDSSSCVACRPPSSKMEV-EEINGIKMKRPTCDICPKACGIGTGSIMSQVTSNIDK 355  
QY 364 FACKKIFGSLARLPESFQDPAASNAPLQPELOVFEETLEETGLYLSAWDLSLPDLS 423  
DB 356 FINTCTINGNLIFLVGIDHPYNALEAIDPELNFRRVRETITGFLINQSWPNNMTDS 415  
QY 424 VFONLAVIRRLIHNAGVSLTLQGLISWLGSLRELSGALLIHNHTHLCFVHVPMD 483  
DB 416 VESNLVYITGRVLYSGLSILLIKQOGTISLQFOSLKEISGNILYITDNSLCTYHINMT 475  
QY 484 QLEFRNPQALLHTANPEDECVEGGLACQOLCARGHCWGPPTQYVNCOSQFLRGQDCEE 543  
DB 476 TLFSTINRIYVIRDNKRAEKCTAEGVNCNHLCSDDCWGPGDQSCRRFSRGRIIES 535  
QY 544 CRYLQGLPREYVVARICLPCHEPCQ-QNGSVYCFEPBADQCYAACHYDPPFCARCS 602

```

Db 536 CNLYDGEFFREPNIGICVBCDPOCEKMEDELLTCHGPGDNCTKCSHFKDGPNVCEKCPD 595
OY 603 GVKPDLSTYPIKFPDEBACOPCPINCHSCVDDDKC-----PAEQ----- 647
Db 596 GIGQANSF--IFRYADPDRECHPCHPNCTOGCNGPTSHDCIYYPMWGHSTLQPHARTPLI 653
OY 648 -----ASPLTSQ----- 654
Db 654 AAGVIGLFLIVGLTFVAVYRRKSIKKKRALRFLTELVELPEPLPSTAPRQAOALRL 713
OY 655 -----NEDLGP----- 660
Db 714 KETELKRVKVLGSGAFVYKGIWPEGETVKIPAIKILNETTGPKANVEKDEALIMA 773
OY 661 ----- 660
Db 774 SMDPHLVRLGCLSPITQLVTLQMPHCCLLEYVHEKDNIGSOLLNMCVOIAKAMX 833
OY 661 ----- 660
Db 834 LEERLVLHNDLAAENVLVKSPNHVKITDFGLARLEGEDEKVENADGKMPKIMALECIH 893
OY 661 ----- 660
Db 894 YRKETHSDVMSGYTITWELMTEFGCKPYDGIPTREIPDLLEKGERLPQPPICTIDVYVM 953
OY 661 -----ASPLDSTFYISLEDDM 678
Db 954 VKCMIDADSRRPKELAEFSRMARDPORYLVIOGDDRMKLPSPNDSKFFQNLDEEDL 1013
OY 679 GDLVDAEEYLVPOQGFCDPAPAGAGVHHNRSSSTFGSDGLTLGLEPSEEARPS- 737
Db 1014 EDMDAEEYLV-QAENIPPP-----ITSTRARIDSNS-----ETGSPRAYTPMSG 1061
OY 738 -----PLAP--SEGASDVFDGDLGMAKAGLSLPTHD 769
Db 1062 NQEVYRIDGEFAAEGQSVYRAPRTSTIPAPVAQAQTATLFPDSCNGTLRKPVAHVE 1121
OY 770 PSPLDRISEDPVPLPS-----ETDGYVAPLTCSPQEVYVNPQVVRQPSPREGRIP 822
Db 1122 DSSTRYSDPTVFAFPERSRGELDEGYTMPMRKPKOEYLNPAVE-----ENPVP 1172
OY 823 AARPGATLERPKTSLPSKNGVVKDVFARFGAVENPEYTLPGGAAPQPPPPA----- 876
Db 1173 SKR-----KNGDLQ-----ALDNPETYNASNG-----PKAEDEYVN 1204
OY 877 -----ESPADFNLVYWDODPPERGA--PSTFKGPT 906
Db 1205 EPLYLNTFANTLGKAEYLLKNITLSMPEKAKAFDNDVNNHSLPRSTLQHPDYLOEYST 1264
OY 907 -----AENPEYL 913
Db 1265 KYFYKONGRIRPIVAENPEYL 1285

```

```

RA Marchionni M.A., Kelly R.A.;
RT "Neuregulins promote survival and growth of cardiac myocytes."
RT Persistence of ErbB2 and ErbB4 expression in neonatal and adult
RT ventricular myocytes."
RL J. Biol. Chem. 273:10261-10269(1998).
RN [2]
RP SEQUENCE OF 848-901 FROM N.A.
RC TISSUE-sciatic nerve;
RX MEDLINE-9122560; PubMed-2025425;
RA Lai C., Lemke G.;
RT "An extended family of protein-tyrosine kinase genes differentially
RT expressed in the vertebrate nervous system."
RL Neuron 6:691-704(1991).
RN [3]
RP SEQUENCE OF 1031-1198 FROM N.A.
RC STRAIN-Sprague-Dawley; TISSUE-Spinal cord;
RX MEDLINE-97184212; PubMed-9030624;
RA Carroll S.L., Miller M.L., Frohert P.W., Kim S.S., Corbett J.A.;
RT "Expression of neuregulins and their putative receptors, ErbB2 and
RT ErbB3, is induced during Wallerian degeneration."
RL J. Neurosci. 17:1642-1659(1997).
CC -1- FUNCTION: SPECIFICALLY BINDS AND IS ACTIVATED BY NEUREGULINS, NRG-
CC 2, NRG-3, HEPARIN-BINDING EGF-LIKE GROWTH FACTOR, BETACELLULIN AND
CC NTAK. INTERACTION WITH THESE FACTORS INDUCES CELL DIFFERENTIATION.
CC NOT ACTIVATED BY EGF, TGF- $\alpha$ , AND AMPHIREGULIN (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- SUBUNIT: HOMODIMER OR HETERODIMER WITH EACH OF THE OTHER ERBB
CC RECEPTORS (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN THE DEVELOPING
CC NERVOUS SYSTEM. EXHIBITS DISTINCT AND HIGHLY REGIONALIZED PATTERNS
CC OF EXPRESSION IN THE ADULT BRAIN, WHERE IT IS MAINLY FOUND IN THE
CC RETICULAR NUCLEUS OF THE THALAMUS. VERY LOW LEVELS IN KIDNEY, AND
CC HEART.
CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE
CC RESIDUES (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb-sdb.ch).
CC -----
DR EMBL: AF041838; AAC08899.1; -.
DR EMBL: U52531; AAC53051.1; -.
DR HSSP: P11362; 1RKG.
DR InterPro: IPR000494; EGFR_L_domain.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002174; Furin-1like.
DR InterPro: IPR001245; Tyr_pkinase.
DR InterPro: IPR004019; YLP_motif.
DR Pfam: PF00757; Furin-1like; 1.
DR Pfam: PF00069; pkinase; 1.
DR Pfam: PF01030; Recep_L_domain; 2.
DR Pfam: PF02757; YLP; 2.
DR PRINTS: PR00109; TYRKINASE.
DR PRODOM: PD000001; Euk_pkinase; 1.
DR SMART: SM00261; FU; 4.
DR SMART: SM00219; Tyrc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;
KW Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
FT SIGNAL 1 25
FT CHAIN 26 1308 RECEPTOR PROTEIN-TYROSINE KINASE ERBB-4.
FT DOMAIN 26 651 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 652 675 POTENTIAL.
FT DOMAIN 676 1308 CYTOPLASMIC (POTENTIAL).

```

```
FT DOMAIN 186 334 CYS-RICH.
FT DOMAIN 496 633 CYS-RICH.
FT DOMAIN 718 985 PROTEIN KINASE.
FT NP_BIND 724 732 ATP (BY SIMILARITY).
FT BINDING 751 751 ATP (BY SIMILARITY).
FT ACT_SITE 843 843 BY SIMILARITY.
FT DISULFID 189 197 BY SIMILARITY.
FT DISULFID 193 205 BY SIMILARITY.
FT DISULFID 213 221 BY SIMILARITY.
FT DISULFID 217 229 BY SIMILARITY.
FT DISULFID 230 238 BY SIMILARITY.
FT DISULFID 234 246 BY SIMILARITY.
FT DISULFID 249 258 BY SIMILARITY.
FT DISULFID 262 289 BY SIMILARITY.
FT DISULFID 293 304 BY SIMILARITY.
FT DISULFID 308 323 BY SIMILARITY.
FT DISULFID 326 330 BY SIMILARITY.
FT DISULFID 503 512 BY SIMILARITY.
FT DISULFID 507 520 BY SIMILARITY.
FT DISULFID 523 532 BY SIMILARITY.
FT DISULFID 536 552 BY SIMILARITY.
FT DISULFID 555 569 BY SIMILARITY.
FT DISULFID 559 577 BY SIMILARITY.
FT DISULFID 580 589 BY SIMILARITY.
FT DISULFID 593 614 BY SIMILARITY.
FT DISULFID 617 625 BY SIMILARITY.
FT DISULFID 621 633 BY SIMILARITY.
FT MOD_RES 1162 1162 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 1188 1188 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 1258 1258 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 1284 1284 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 253 253 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 358 358 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 410 410 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 473 473 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 495 495 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 548 548 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 576 576 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 620 620 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 1062 1062 S -> N (IN REF. 3).
FT CONFLICT 1080 1082 PYT -> SYR (IN REF. 3).
SQ SEQUENCE 1308 AA; 146957 MW; D944B0996A0B41 CRC64;
```

Query Match 29.9%; Score 1519; DB 1; Length 1308;  
Best Local Similarity 28.1%; Pred. No. 2e-80;  
Matches 371; Conservative 143; Mismatches 364; Indels 442; Gaps 25;

```
QY 1 MELA-ALCRMGILL-ALLPFGASTOYCTGDMKRLRLPASPTHLDMLRHLXGCGQVQ 57
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MKLATGLMWGSLVLAARTVQPSASQSCAGTENKLSLSDLEOYRALRKYYENCEVVM 60

QY 58 GNEFLYLPINASLSTLQDIOEVQVILAHNOVROPRLRLRYKGTQGFEDNYALVL 117
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 GNEITISIEHRDLSFLRSIREVYVVALNCPRYPLENLRIRKTKLEDEYALAIIF 120

QY 118 DNGDPLNNTTPVGTGASPGGRLQLRSITELKGGVILQORNPOLCYOTILMKDIEHNN 177
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 LNYKKDGNF-----GQDELGLKNLTLELNGGVYVDOKKLCYALTJHWDIVRNPW 171

QY 178 QALTLTIDYNRACHPCSPWCKSGRCMGESSEDCQSLTRTVCAAGC-ARCKGLPTDCC 236
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 172 PSNMTLVSTIGSSGCGRHKSCGTG-RCKMGPTEHNCQILTRIVCAEDCGRGRTGYVSDCC 230

QY 237 HEQCAAGCTGPKHSDCLACLFHNSGICEELCPALVTYNTDFESMPNDEGRYTFGASCV 296
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 231 HRECAAGCGSGPKDPCFACMNFNDSGACVQCPTQFVYNPPTFQLEHNFNAKYTGACVY 290

QY 297 TACGYNTLSTDVGSCITVCPAHNEVTAEDCTQCEKSKCARVYCGLGHEHLEVRAY 356
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 291 KCPCHNFV-VDSSSCVACBPSSKMEV-EENGIKMKCKPCTDIPRACDGIIGSLMSQTV 348
```

```
QY 357 TSANIOEPAGCKIIFGSLAFLESFDDPASNTAPLQPEQLQVEETLEITGYLISAMP 416
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 349 DSSNIDKRIKNTKINGNIIFLVTHGIDPYNALDAIDPEKLNVRFTYREITGFINIQWP 408

QY 417 DSLPDLVFOULQYIRGILHNGAVSLTLOGLGISWGLRSRLRELGSGLAIHHNTHICE 476
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 409 PNMDFESFVSLVYIGGVVLSGLSLILKQOGITSLQFOSIKETISAGNITYTNSNICY 468

QY 477 VHTVPMQDLFNPHQALHTNPDEDECVSGELCHOLCARGHCGMGPTQCVNCSQFLR 536
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 469 YHTNMVTLTLESTVQRIYTRDNRRAENCTABGMVCNHLCSNDGCMGEPDQCLCRFRSR 528

QY 537 GOECVEECRYLOGLPREXVNAHCLPCHRECP-ONGSVTCGPEADQVCAAHYKDPPE 595
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 529 GKICIESCNLYDGEFERENGISICEQDSQCEKMEKEDGLITGHGPRDCTGCSHPKQPN 588

QY 596 CVARCPGCVKPDLSYMPWKFPDEBACQPCPINCSTHSCVDLDKGC-----PA 644
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 589 CVERCPDVLQAGNSF--TFKYADODRECHPCHPNCTQCGNSPTSHDCIYYPWTGHTLPQ 646

QY 645 EQP-----ASPLTSQ----- 654
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 647 HAPRPLIAGVIGGLFVLYMALTFYAVYVRKKSIAKKALRRLFETELVEPLTSGTAPN 706

QY 655 -----NEDLGP----- 660
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 707 QALRLIKETELKRVKVLGSGAFGVYKGIWVPEGETYKIPVAKIKLNETGPRANVEFM 766

QY 661 ----- 660
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 767 DEALIMASVDHPLHVRLLGVLCSPTIOLVTLQMPHGCLLEVYHBEKDNIGSOLLNMCVQ 826

QY 661 ----- 660
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 827 IAKGMVYEEPRVLYHRDLAANVLYKSPNHKKTIDFGIARLLEDEKEYNADGSKMPLX 886

QY 661 ----- 660
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 887 MALBCIHYRFTHQSDVMSYGVTIWELMTFGGKRYDGIPTREIPDLLEKGRPLPQPICT 946

QY 661 -----ASPLDSTFYRS 671
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 947 IDVYIVMKWMIDADSRRPKRELAERSMARDPQRYLYTIGDDRMKLPSPNDSKFEON 1006

QY 672 LLEDMDGDLVDAEYLYVPOOGFPCDP-----APGAGGVH 708
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1007 LLEDDELDMDAEBEYLYP-QAFNIPPTIYSRRIDSNRSEIGHSPRPATTPSGSGFV 1065

QY 709 HHRSSSTRSGGDLTGLLEPSEEBAPRSLAPSEAGSDVFDGLGGAAGLQSLPTH 768
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1066 YODGGFATQOG--MPMYTATTSIPEAPYA--QGAATAEMFDDSCNGTLRKFPVPHVQ 1120

QY 769 DSPFLQRTSEDPYVPLBS-----ETDGYAPLTCSPQPYVNOQDVRPQPPRQSGPL 821
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1121 EDSSTQRSAPPTVAFABERNRAELDEGYMTPHMDPKOBYLPLVVENPEVPSRRKGD 1180

QY 822 PAAPRAGATLEPRK-----TSPGKNGVKYDVE-----AFGAVENPEYLTPOGGAAP 869
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1181 QA-----LDNPEYHSASSPPK---AEDYVNEPLYLNTFNALGNAETKMSLSL 1230

QY 870 QHPHPAPSPAFDNLVYWDODPERGA--PSTKGTPT-----AENPEYL 913
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1231 E-----KAKKAFDNPDMVNNHSLPPRSTLQHPDYLDQESTKTFYKQNGKIRPIVAENPEYL 1285
```

RESULT 8  
EGFR\_CHICK  
ID EGFR\_CHICK STANDARD; PRT; 703 AA.  
AC P13387;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Epidermal growth factor precursor (EC 2.7.1.112) (CER)

[illegible]

DE (c-erbB3) (tyrosine kinase-type cell surface receptor HER3).  
GN ERBB3 OR HER3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90083234; PubMed=2687875;  
RA Kraus M.H., Issing W., Miki T., Popescu N.C., Aaronson S.A.;  
RT "Isolation and characterization of ERBB3, a third member of the  
RT ERBB/epidermal growth factor receptor family: evidence for  
RT overexpression in a subset of human mammary tumors.";  
RL Proc. Natl. Acad. Sci. U.S.A. 86:9193-9197(1989).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90311312; PubMed=2164210;  
RA Plioman G.D., Whitley G.S., Neubauer M.G., Green J.M., McDonald V.L.,  
RA Todaro G.J., Shoyab M.;  
RT "Molecular cloning and expression of an additional epidermal growth  
RT factor receptor-related gene.";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:4905-4909(1990).  
RN [3]  
RP SEQUENCE FROM N.A. (SHORT FORM).  
RX TISSUE=Placenta; PubMed=7685162;  
RA Katoh M., Yazaki Y., Sugimura T., Terada M.;  
RT "c-erbB3 gene encodes secreted as well as transmembrane receptor  
RT tyrosine kinase.";  
RL Biochem. Biophys. Res. Commun. 192:1189-1197(1993).  
CC -1- FUNCTION: BINDS AND IS ACTIVATED BY NEURULIN AND NTRK.  
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein  
CC tyrosine phosphate.  
CC -1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS  
CC (POTENTIAL).  
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (LONG FORM) AND  
CC SECRETED (SHORT FORM).  
CC -1- ALTERNATIVE PRODUCTS: TWO FORMS, A LONG TRANSMEMBRANE FORM AND A  
CC SHORT SECRETED FORM, EXIST DUE TO ALTERNATIVE SPLICING.  
CC -1- TISSUE SPECIFICITY: EPITHELIAL TISSUES AND BRAIN.  
CC -1- DOMAIN: THE CYTOPLASMIC PART OF THE RECEPTOR MAY INTERACT WITH THE  
CC SH2 OR SH3 DOMAINS OF MANY SIGNAL-TRANSDUCING PROTEINS.  
CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE RESIDUES  
CC AND PROMOTES ITS ASSOCIATION WITH THE P85 SUBUNIT OF  
CC PHOSPHATIDYLINOSITOL 3-KINASE (BY SIMILARITY).  
CC -1- DISEASE: OVEREXPRESSED IN A SUBSET OF HUMAN MAMMARY TUMORS.  
CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: M29366; AAA35790.1; -  
DR EMBL: M34309; AAA35979.1; -  
DR EMBL: S61953; AAB26935.1; -  
DR PIR: A36223; A36223.  
DR HSSP: P1362; 1FGK.  
DR Genew: HGNC:3431; ERBB3.  
DR MIM: 190151; -  
DR InterPro: IPR000494; EGFR\_L\_domain.  
DR InterPro: IPR000719; Euk\_pkinase.  
DR InterPro: IPR001245; Tyr\_pkinase.  
DR Pfam: PF00069; pkinase; 1.  
DR Pfam: PF00757; Furin-like; 1.  
DR Pfam: PF01030; Recep\_L\_domain; 2.  
DR ProDom: PD000001; Euk\_pkinase; 1.  
DR SMART: SM00261; FU; 3.  
DR SMART: SM00219; TyrKc; 1.

DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; FALSE\_NEG.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; FALSE\_NEG.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;  
KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;  
KW Alternative splicing.  
FT SIGNAL 1  
FT CHAIN 19  
FT DOMAIN 20 1342  
FT TRANSMEM 20 643  
FT DOMAIN 644 664  
FT DOMAIN 665 1342  
FT DOMAIN 709 966  
FT NP\_BIND 715 723  
FT BINDING 742 742  
FT ACT\_SITE 834 834  
FT DISULFID 186 194  
FT DISULFID 190 202  
FT DISULFID 210 218  
FT DISULFID 214 226  
FT DISULFID 227 235  
FT DISULFID 231 243  
FT DISULFID 246 255  
FT DISULFID 259 286  
FT DISULFID 290 301  
FT DISULFID 305 320  
FT DISULFID 323 327  
FT DISULFID 500 509  
FT DISULFID 504 517  
FT DISULFID 520 529  
FT DISULFID 533 549  
FT DISULFID 552 565  
FT DISULFID 556 573  
FT DISULFID 576 585  
FT DISULFID 589 610  
FT DISULFID 613 621  
FT DISULFID 617 629  
FT CARBOHYD 126 126  
FT CARBOHYD 250 250  
FT CARBOHYD 353 353  
FT CARBOHYD 408 408  
FT CARBOHYD 414 414  
FT CARBOHYD 437 437  
FT CARBOHYD 469 469  
FT CARBOHYD 522 522  
FT CARBOHYD 566 566  
FT CARBOHYD 616 616  
FT VARSPLIC 141 183  
FT VARSPLIC 184 1342  
FT CONFLICT 560 560  
FT CONFLICT 1064 1064  
SQ SEQUENCE 1342 AA; 148097 MW; 7201E7F66CA374BD CRC64;  
Query Match 28.3%; Score 1435; DB 1; Length 1342;  
Best Local Similarity 43.6%; Pred. No. 1; 4e-75;  
Matches 278; Conservative 92; Mismatches 244; Indels 24; Gaps 12;  
QY 10 GLLALALPPGAA--STQVCTGIDMKRLRPASEPETHLDMRLHYOGGVQVGMELTYLPT 67  
DB 11 GLIFSLARGSEVGNSSQAVCGTLNGLSVGDENQYOTLYKLYEREVEVVMGLIEVLTGH 70  
QY 68 NASLSFLDIOIQEVGVYLLAHNOVROVPLRLRIYRGTLQFDENVALAVLNGDDLLNNT 127  
DB 71 MADLSLQWIREVTGTYVLAAMEFSTLPPLNLRVVRGVYDGKRALEWMA---LNTYT 125  
QY 128 PVTGASPGGLRLRLQSLRSLTEILKGGVLIQRNPOLCYODTILMKDIFHKNKQALATLIDTN 187  
DB 126 ----NSSHALRQLRLQLTEILISGGVYTEKNKLCMHMDIDRDIVRDRD---AIIYVKD 178  
QY 188 RSRACHPCSPMKSGRCMGESSEDCQSLIRYVACAGC--ARCKGRLPTDCHQCAAGCTG 246  
DB 179 NGRSCPCHGVCKG--RCWGPGESEDCQTLTKTICAPCGNCHGPGPNOCCHDECAAGCGSG 237

247 PKHSDCLACIHPNHSIGCEHCALPATYNTDFESMPNPEGRYTFGASCTACPYNTLST 306  
DB 238 P0DIDCFACHFNFDSCACVRCPOPLVYNTLTFOLEPNPHTKYOGGCVASCPHNFV-V 296  
OY 307 DVSCSTVCPVCLHNOETVAEGRGTCORCEKSPCARVCGYGLMEHLREVRATSNIOEFAG 366  
DB 297 DQNSCRACPRDKMEVD-KKGLMKCEPCGGCLCKRACEGTSG---SRQYVDDSSNDIGFVN 353  
OY 367 CKRIFFSLAFPEPSFDGDPASNTAPLOEQLQVETLEETGYLYISAWPDSLPDISVFO 426  
DB 354 CTKILGNLDELITGLNDDPMHKIPALDPEKLVNFRVETITGYLNIQSMPPHNFNSVFS 413  
OY 427 NLQVINGRIIHNAGAYS-LTLOGIGISMLGRLSRLSELGSLALLHNHTLHCFVITVPDOL 465  
DB 414 NLVTIGRSLYNGFSLLIMKNLNTVSLGFRSLKEISAGRIYISANQLCYHNSLMTKY 473  
OY 486 FRAPHOALLHTA-NRPEDECVGEGELACHOHLCARGHCMPGPTOCVNCOSFLRGOCVEEC 544  
DB 474 LRQPTERRLDIKHNRPRRCVAGKVCDDPLCSSGGCGPFGCGLSCRNTSRGVCYTHC 533  
OY 545 RVLGSLPREVYNARHCLPCHPECOPOGNSVTCFGEPAODCVACAHYKDPPECVAPRSGV 604  
DB 534 NFLNGEPREFAHBAECFSCHEPCOPEMGATKNGSGSDTCAOCACAFHRDFGPHCVSSCPHGV 593  
OY 605 KPDLSTMPIMKFPDEGACQPCPINCITHSCVDLDDKGC 642  
DB 594 LG--ANGPIYKYPDVONECRCHENCTQCGKPELDC 629  
RESULT 10  
ERB3\_RAT  
ID ERB3\_RAT STANDARD; PRT; 1339 AA.  
AC 062759; 062955;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Receptor protein-tyrosine kinase erbB-3 precursor (EC 2.7.1.112)  
DE (C-erbB3).  
GN ERB3.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Rattus.  
OX NCBI\_Taxid=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sprague-Dawley; TISSUE=Liver;  
RA MEDLINE=96096535; PubMed=8522190;  
RA Hellyer N.J., Kim H.-H., Greaves C.H., Sierke S.L., Koland J.G.;  
RT "Cloning of the rat ErbB3 cDNA and characterization of the  
RT recombinant protein";  
RL Gene 165:279-284(1995).  
RN [2]  
RP REVISIONS TO 85: 513 AND 565.  
RA Hellyer N.J., Koland J.G.;  
RL Submitted (DEC-2001) to the EMBL/Genbank/DBJ databases.  
RN [3]  
RP SEQUENCE OF 922-1097 FROM N.A.  
RC STRAIN=Sprague-Dawley; TISSUE=Sciatic nerve;  
RA MEDLINE=97184212; PubMed=9030624;  
RA Carroll S.L., Miller M.L., Frohner P.W., Kim S.S., Corbett J.A.;  
RT "Expression of neuroligin and their putative receptors, ErbB3 and  
RT ErbB4, is induced during Wallerian degeneration.";  
RL J. Neurosci. 17:1642-1653(1997).  
CC -1- FUNCTION: BINDS AND IS ACTIVATED BY NEUREGULINS AND NTRK.  
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein  
CC tyrosine phosphate.  
CC -1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS  
CC (POTENTIAL).  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- DOMAIN: THE CYTOPLASMIC PART OF THE RECEPTOR MAY INTERACT WITH THE  
CC SH2 OR SH3 DOMAINS OF MANY SIGNAL-TRANSDUCING PROTEINS.  
CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE RESIDUES

CC AND PROMOTES ITS ASSOCIATION WITH THE P85 SUBUNIT OF  
CC PHOSPHATIDYLINOSITOL 3-KINASE.  
CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: U29339; AAC28498.2; -;  
DR EMBL: U52530; AAC33050.1; -;  
DR HSSP: P11362; 1FGK.  
DR InterPro: IPR000494; EGFR\_L\_domain.  
DR InterPro: IPR000719; Euk\_kinase.  
DR InterPro: IPR002174; Furin-1-like.  
DR InterPro: IPR001245; Tyr\_kinase.  
DR Pfam: PF00069; pkinase; 1.  
DR Pfam: PF00757; Furin-1-like; 1.  
DR Pfam: PF01030; Recep\_L\_domain; 2.  
DR PRINTS: PR00109; TYRKINASE.  
DR ProDom: PD000001; Euk\_pkinase; 1.  
DR SMART: SM00261; RU: 5  
DR SMART: SM00219; TYRK; 1.  
DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; FALSE\_NEG.  
DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR Transmembrane: Glycoprotein; Multigene family; Receptor; Signal;  
KW Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.  
FT SIGNAL 1 19  
FT CHAIN 20 1339  
FT DOMAIN 20 643  
FT TRANSMEM 644 662  
FT DOMAIN 663 1339  
FT DOMAIN 183 259  
FT DOMAIN 707 964  
FT NP\_BIND 713 721  
FT BINDING 740 740  
FT ACT\_SITE 832 832  
FT DISULFID 186 194  
FT DISULFID 190 202  
FT DISULFID 210 218  
FT DISULFID 214 226  
FT DISULFID 227 235  
FT DISULFID 231 243  
FT DISULFID 246 255  
FT DISULFID 259 286  
FT DISULFID 290 301  
FT DISULFID 305 320  
FT DISULFID 323 327  
FT DISULFID 500 509  
FT DISULFID 504 517  
FT DISULFID 520 529  
FT DISULFID 533 549  
FT DISULFID 537 573  
FT DISULFID 566 585  
FT DISULFID 576 610  
FT DISULFID 589 621  
FT DISULFID 613 629  
FT CARBOHYD 126 126  
FT CARBOHYD 250 250  
FT CARBOHYD 353 353  
FT CARBOHYD 408 408  
FT CARBOHYD 414 414  
FT CARBOHYD 437 437  
FT CARBOHYD 469 469  
FT CARBOHYD 522 522  
FT CARBOHYD 566 566  
FT CARBOHYD 616 616  
FT CONFLICT 1028 1028  
SQ SEQUENCE 1339 AA; 147545 MW; 0AA5F2402BBDFE CRC64;









RA Zhong X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gbbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 RN [17]  
 RP SEQUENCE OF 959-1078 FROM N.A.  
 RC STRAIN-Daekwanryeong;  
 RX MEDLINE-85137938; PubMed-2983232;  
 RA Wadsworth S.C., Vincent W.S. II, Bilocheau-Wentworth D.;  
 RT "A *Drosophila* genomic sequence with homology to human epidermal  
 RT growth factor receptor.";  
 RL Nature 314:178-180(1985).  
 RN [18]  
 RP SEQUENCE OF 1133-1137; 1155-1167 AND 1209-1216 FROM N.A., AND MUTATION  
 RP ANALYSIS.  
 RX MEDLINE-92038942; PubMed-1936959;  
 RA Raz E., Schejter E.D., Shilo B.Z.;  
 RT "Intallelic complementation among DER/flb alleles: implications for  
 RT the mechanism of signal transduction by receptor-tyrosine kinases.";  
 RL Genetics 129:191-201(1991).  
 RN [9]  
 RP REVIEW.  
 RX MEDLINE-97248481; PubMed-9094709;  
 RA Perrimon N., Perkins L.A.;  
 RT "There must be 50 ways to rule the signal: the case of the *Drosophila*  
 RT Egr receptor.";  
 RL Cell 89:13-16(1997).  
 CC -1- FUNCTION: BINDS TO FOUR LIGANDS: SPITZ, GURKEN, VEIN AND ARGOS,  
 CC WHICH IS AN ANTAGONIST. TRANSDUCES THE SIGNAL THROUGH THE RAS-RAF-  
 CC MAPK PATHWAY. INVOLVED IN A MYRIAD OF DEVELOPMENTAL DECISIONS.  
 CC CRITICAL FOR THE PROLIFERATION OF IMAGINAL TISSUES, AND FOR THE  
 CC DETERMINATION OF BOTH THE ANTERO-POSTERIOR AND DORSO-VENTRAL  
 CC POLARITIES OF THE OOCYTE. IN THE EMBRYO, PLAYS A ROLE IN THE  
 CC ESTABLISHMENT OF VENTRAL CELL FATES, MAINTENANCE OF AMNIOSEROSA  
 CC AND VENTRAL NEUROECTODERMAL CELLS, GERM BAND RETRACTION, CELL FATE  
 CC SPECIFICATION IN THE CENTRAL NERVOUS SYSTEM AND PRODUCTION OF  
 CC CUTICLE.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein  
 CC tyrosine phosphate.  
 CC -1- SUBCELLULAR LOCATION: ISOFORMS TYPE I AND TYPE II ARE TYPE I  
 CC MEMBRANE PROTEINS. ISOFORM TYPE III IS A TYPE III MEMBRANE  
 CC PROTEIN.  
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS. TYPE I (SHOWN HERE), TYPE II AND  
 CC TYPE III ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- TISSUE SPECIFICITY: UBQUITOUSLY EXPRESSED IN EMBRYOS. IN LARVAE,  
 CC UNIFORM EXPRESSION IS SEEN IN WING DISKS, GENITAL DISK, ANLAGEN OF  
 CC TESTIS AND OVARY, AND BRAIN CORTEX. IN EYE-ANTENNA DISK, HIGHEST  
 CC EXPRESSION IS ANTERIOR TO MORPHOGENETIC FURROW, LEVELS REMAIN HIGH  
 CC IN PHOTORECEPTOR PRECURSOR CELLS. THIS PATTERN IS REVERSED IN  
 CC POSTERIOR EYE DISK. IN ADULTS EXPRESSION IS HIGH IN BRAIN CORTEX  
 CC AND THORACIC AND ABDOMINAL GANGLIA.  
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: AF052754; AAC08536.1; -  
 DR EMBL: AF052753; AAC08536.1; JOINED.  
 DR EMBL: AF052754; AAC08535.1; -  
 DR EMBL: AF052752; AAC08535.1; JOINED.  
 DR EMBL: K03054; AAA51462.1; -  
 DR EMBL: K03417; AAA51460.1; -  
 DR EMBL: K03416; AAA50965.1; -  
 DR EMBL: K03418; AAA51461.1; -  
 DR EMBL: AF109077; AAD26134.1; -  
 DR EMBL: AF109078; AAD26132.1; -  
 DR EMBL: AF109082; AAD26132.1; JOINED.  
 DR EMBL: AF109082; AAD26133.1; -  
 DR EMBL: AF109078; AAD26133.1; -

DR EMBL: AF109084; AAD26133.1; JOINED.  
 DR EMBL: AF109079; AAD26130.1; -  
 DR EMBL: AF109081; AAD26130.1; JOINED.  
 DR EMBL: AF109079; AAD26131.1; -  
 DR EMBL: AF109083; AAD26131.1; JOINED.  
 DR EMBL: AF109080; AAD26135.1; -  
 DR EMBL: AE003454; AAF46732.1; -  
 DR EMBL: X02293; CAA26157.1; -  
 DR EMBL: X78920; CAA55523.1; -  
 DR EMBL: X78918; CAA55521.1; -  
 DR EMBL: X78919; CAA55522.1; -  
 DR PIR: A00640; GOFEE.  
 DR HSP; P11362; IFEK.  
 DR Flybase: FBgn0003731; Egr.  
 DR InterPro: IPR000494; EGFR\_L.domain.  
 DR InterPro: IPR000719; Euk\_kinase.  
 DR InterPro: IPR002174; Furin-like.  
 DR InterPro: IPR001245; Tyr\_kinase.  
 DR Pfam: PF00069; Kinase; 1.  
 DR Pfam: PF00757; Furin-like; 1.  
 DR Pfam: PF01030; Recep\_L.domain; 2.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR ProDom: PD000001; Euk\_kinase; 1.  
 DR SMART: SM00261; FU; 7.  
 DR SMART: SM00219; Tyrc; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR Transmembrane: Glycoprotein; Receptor; Phosphorylation; Transferase;  
 DR Tyrosine-protein kinase; ATP-binding; Signal; Alternative splicing;  
 DR Developmental protein.  
 FT CHAIN 1 1426  
 FT DOMAIN 31 868  
 FT TRANSSEM 869 889  
 FT DOMAIN 890 1426  
 FT DOMAIN 938 1196  
 FT NP\_BIND 944 952  
 FT BINDING 971 971  
 FT ACT\_SITE 1063 1063  
 FT MOD\_RES 902 902  
 FT PHOSPHORYLATION (BY PKC) (BY SIMILARITY).  
 FT POTENTIAL.  
 FT EPIDERMAL GROWTH FACTOR RECEPTOR.  
 FT EXTRACELLULAR (POTENTIAL).  
 FT POTENTIAL.  
 FT CYTOPLASMIC (POTENTIAL).  
 FT PROTEIN KINASE.  
 FT ATP (BY SIMILARITY).  
 FT ATP (BY SIMILARITY).  
 FT BY SIMILARITY.  
 FT PHOSPHORYLATION (BY PKC) (BY SIMILARITY).  
 FT  
 Query Match 20.0%; Score 1013.5; DB 1; Length 1426;  
 Best Local Similarity 33.0%; Pred. No. 3.3e-51;  
 Matches 231; Conservative 93; Mismatches 280; Indels 97; Gaps 20;  
 Oy 24 OVCSTDMKRLRPASPERHLMRLTYOGCOVGNLETLPT-NASTSLFDIOEVOG 82  
 Db 100 KICICTKRLSVSPSKKEHHYRLNDRYNTCTYVDGNLKLTPDNDLDSFLDIREVTG 159  
 Oy 83 YVLAHNOVROVPLQRLRIVRGTOLE-----EDNYALAVLDNGDPLNNTTPTVTGASPGGL 137  
 Db 160 YILSHVDKVKVPPKLIQINGRTLFLSVVEEKALV-----TSKM 203  
 Oy 138 RELQIRSLTEILKGVLIQIRNPOLCYODTILMKDIFHKNNOLALTLDTNSRACHPCSP 197  
 Db 204 YLLEIPDLRDVLDNGVGHNNYINLCMHMTIOWSELVSNGTDAYYNYDTAPERCPCKE 263  
 Oy 198 MCKGSRGCESESDQSLRTVCAAGCA--CCKGPLPDCCHCCAGCAGCTPKKSSDCLAC 255  
 Db 264 SCYHS-CWGEGRKCKQKSKLTCSPOCAGRCYGGKPRECHLPCAGGCTGPTPKDCLAC 322  
 Oy 256 LHFNSGICELHCPALVYNTDTEFSPNPEGRYTFGASCYTACPYNTSLDVGSCITVC 315  
 Db 323 KNFPENAVSKKECPMRKYNFTTYLLENPBGKAYGATCYKECP-GHLLDNGACVASC 381  
 Oy 316 PLHNOVTAEDGTORCEKSCPCARVCYGLGMEHLREYRAVTSANIOEFACCKKIIFGSLA 375  
 Db 382 PODKMDKGE-----CVCNCPCEPCPTCGVYLH-----AGNIDSFRCVTYDGNIR 428  
 Oy 376 FLPESEFDG--DPASNTA-----PLQPBQLOVFELEITGYIYISAMPDPLPLSVYQN 427  
 Db 429 ILDTFSGFDVYANYTMGPRIYPLDPERREVFTVKEITGYIYLNIEGTHQPRNLSYFRN 488



FT DISULFID 670 693 POTENTIAL.  
 FT DISULFID 696 703 POTENTIAL.  
 FT DISULFID 700 715 POTENTIAL.  
 FT DISULFID 717 731 POTENTIAL.  
 FT DISULFID 735 750 POTENTIAL.  
 FT DISULFID 753 763 POTENTIAL.  
 FT DISULFID 757 771 POTENTIAL.  
 FT DISULFID 774 787 POTENTIAL.  
 FT DISULFID 791 805 POTENTIAL.  
 FT CARBOHYD 91 91 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 169 169 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 376 376 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 561 561 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 655 655 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 746 746 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 776 776 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARIANT 368 368 C -> Y (IN SY10).  
 FT VARIANT 469 469 C -> R (IN MN216).  
 FT VARIANT 700 700 C -> W (IN MN23).  
 FT VARIANT 753 753 C -> Y (IN SY11).  
 FT VARIANT 1065 1065 T -> I (IN SY16).  
 FT VARIANT 1074 1074 G -> E (IN SY7).  
 SQ SEQUENCE 1323 AA: 150510 MW: 680307E53BEF99 CRC64;

Query Match 14.1%; Score 718.5; DB 1; Length 1323;  
 Best Local Similarity 28.0%; Pred. No. 3.2e-34;  
 Matches 191; Conservative 107; Mismatches 256; Indels 129; Gaps 28;

OY . 25 VCTGDMKRLRASPETHLMDRLHLYGCGVQVQGNLETLPTN----- 68  
 Db . 39 LCGTGNGISRYGTNT-LEDELTMYRGCRRYGNYLLETWLEANIKKWRSTNSTVDPK 97  
 OY 69 -----ASLSFLQDIOEVGYVLLAHNOVROVPLRLHVRGTQLEEDNYALAVLNDGDP 122  
 Db 98 NEDSPKSLNFEEDNEELRGSLLIYRANIKRSFRLRYGVGEVFNH-ALYIKHNK- 155  
 OY 123 LNNTPVTGASBGGLRELQLSLETLKGVLIQNPOLCY-QDTILMKDIFHKNNQAL 181  
 Db 156 -----VHEVVMRELRYIRNGSVTIQDNPKMYIGDKIDMKELLYDPD--VQ 199  
 OY 182 TLIDNRSRACH-----PCSPMKGSRGSESEDCOSLTRVCAGGAGC---KGPL 231  
 Db 200 KVEITNSHOCYONGKSNMAKCHESC-NDKCMWSGNDCCORYRSVCPKSCSCFYSNSTS 258  
 OY 232 PTDCHEOCAGCTGPKHSDCLACLFHNSGICELHCPALVYVNDTFESMPNPGRYTF 291  
 Db 259 SYECODSACLGCGTGHGPKNCKIACKYELDGCITCPSRKIFNFKTGLVFNPDGRYQN 318  
 OY 292 GASCTACPYNYL-STDVGSCTLYC-PLHNOEVTAEDESTORCEK-SKPCARVCYGLGME 348  
 Db 319 GNHCYKCEPPELLIENDV--CVRHCSGDGHYDATKD--VRECEKCRSSSCPRTCVDG-- 372  
 OY 349 HLREBANTSANIOEFACKKIFGSLAFLPESFODDPAASNAPLOPELOVFEETIELTG 408  
 Db 373 HL-----TNETLKNLECEQIDGHL-ILIEAF-----TYEOLKYLETVKIYSE 414  
 OY 409 YLYISAMPDLSPLSVFONLOVIRGRILHNGAVSLTL-OGGISWLGKRSJRELSSGAL 467  
 Db 415 YITIV--QONFYDLKFLKNLOITIEGRKLIHNYRMALAIYQCDDELESLNLSKLITGAVL 472  
 OY 468 IHNHTLFCFTVPMDOLE-----RNPQOALLHTANRPDECVGEGELACHQCLARGCHGPR 523  
 Db 473 IMKNHRLCYVSKIDMSIITSKGNKNSPLAIAENRSDKLETEBGRVGDKNKNGCKGK 532  
 OY 524 GPLOCVNCSQLRGQECYEECRVLOGLPREYVNAHRCPLCHPEOCQPOGNSVTCFGEPRADQ 583  
 Db 533 EPECLCEKTKWSVGTCEKCDT-KGFLRNQTSMK-CERCEPECE-----TONGELGELD 584  
 OY 584 CVACAH-----YKDPFCVARGCPSSGVKRPDLSTYMPIMKPFDEBACQPCPINCINCHSCVD 636  
 Db 585 CLTGRHKLTIYNSDFGNRMCEVHDC-----PVSHFPQKNKVCERKCHPTCY----- 628

OY 637 LDDKGPAPQRASPLTQSNEDLG 659  
 Db 629 --DNCG-----TGPDSNLG 640  
 RESULT 14  
 ID ERB2\_MOUSE STANDARD; PRT; 245 AA.  
 AC P70424; 061525;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Receptor protein-tyrosine kinase erbB-2 (EC 2.7.1.112) [p185erbB2]  
 DE (NEU proto-oncogene) (C-erbB-2) (Fragments).  
 GN ERBB2 OR NEU.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN (1)  
 RP SEQUENCE OF 1-149 FROM N.A.  
 RP STRAIN-CD-1; TISSUE=uterus;  
 RC MEDLINE=97200814; PubMed=9048643;  
 RX Lin J., Day S.K., Das S.K.;  
 RA "Differential expression of the erbB2 gene in the perimplantation  
 RT mouse uterus: potential mediator of signaling by epidermal growth  
 RT factor-like growth factors.";  
 RL Endocrinology 138:1328-1337(1997).  
 RN [2]  
 RP SEQUENCE OF 150-245 FROM N.A.  
 RP MEDLINE=96069911; PubMed=7589796;  
 RX Moscoso L.M., Chu G.C., Gautam M., Noakes P.G., Merlie J.P.,  
 RA Sanes J.R.;  
 RT "Synapse-associated expression of an acetylcholine receptor-inducing  
 RT protein, ARIA/heretulin, and its putative receptors, ErbB2 and ErbB3,  
 RT in developing mammalian muscle.";  
 RL Dev. Biol. 172:158-169(1995).  
 CC -1- FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX,  
 CC ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GP30 IS A  
 CC POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-  
 CC ALPHA AND AMPHIREGULIN.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein  
 CC tyrosine phosphate.  
 CC -1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS  
 CC (POTENTIAL).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN UTERINE EPITHELIAL  
 CC CELLS. IN THE MUSCLE, EXPRESSION LOCALIZES TO THE SYNAPTIC SITES  
 CC OF MUSCLE FIBERS.  
 CC -1- DEVELOPMENTAL STAGE: ON DAYS 1-4 OF PREGNANCY, ERBB2 IS DETECTED  
 CC PRIMARILY IN EPITHELIAL CELLS, THE DAY 1 UTERUS SHOWING THE  
 CC HIGHEST ACCUMULATION. ON DAY 5, THE EPITHELIUM AND THE  
 CC DECIDUALIZING SPROMAL CELLS AROUND THE IMPLANTING BLASTOCYST  
 CC EXHIBIT ACCUMULATION OF THIS RECEPTOR. ON DAYS 6-8, THE EXPRESSION  
 CC PERSISTS IN THE EPITHELIUM AT BOTH THE IMPLANTATION AND  
 CC INTERIMPLANTATION SITES IN ADDITION TO MODEST LEVELS IN THE  
 CC SECONDARY DECIDUAL ZONE. ON DAYS 7 AND 8, ACCUMULATION IS ALSO  
 CC PROMINENT IN THE TROPHOBLASTIC GIANT CELLS.  
 CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE  
 CC RESIDUES (BY STIMULARTY).  
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: U71126; AAB17380.1; -;  
 CC DR EMBL: L47239; AAA93532.1; -;  
 CC DR HSSP: P11362; IFGK.





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2003, 14:44:10 ; Search time 9.60392 Seconds  
(without alignments)  
3074.904 Million cell updates/sec

Title: US-09-854-356-7  
Perfect score: 3954  
Sequence: 1 MELALCRWGLLALLRPGA.....GFCPPDPAPGAGVHHRHR 712

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3776	95.5	1255	1	ERB2_HUMAN
2	3205	81.1	1257	1	ERB2_RAT
3	3185	80.6	1254	1	ERB2_MESAU
4	1533.5	38.8	1210	1	EGFR_MOUSE
5	1532	38.7	1210	1	EGFR_HUMAN
6	1509.5	38.2	703	1	EGFR_CHICK
7	1451	36.7	1308	1	ERB4_HUMAN
8	1440.5	36.4	1308	1	ERB4_RAT
9	1435	36.3	1342	1	ERB3_HUMAN
10	1374.5	34.8	1339	1	ERB3_RAT
11	1330	33.6	1167	1	XMRK_XIPMA
12	1013.5	25.6	1426	1	EGFR_DROME
13	718.5	18.2	1323	1	LY23_CAEBL
14	384	9.7	1363	1	ILPR_BRALA
15	383	9.7	1300	1	IRR_MOUSE
16	371	9.4	1300	1	IRR_CAVPO
17	368	9.3	1382	1	INSR_HUMAN
18	364	9.2	581	1	IRR_RAT
19	362.5	9.2	1383	1	INSR_RAT
20	361	9.1	1372	1	INSR_MOUSE
21	358	9.1	1297	1	IRR_HUMAN
22	353	8.9	1477	1	HTK7_HYDDE
23	326	8.2	1390	1	INSR_MOUSE
24	324	8.2	1607	1	MYR_LYMST
25	316	8.0	1370	1	IGIR_RAT
26	314	7.9	1373	1	IGIR_MOUSE
27	311.5	7.9	1367	1	IGIR_HUMAN
28	307	7.8	1246	1	INSR_DROME
29	271.5	6.9	1696	1	PKC5_MOUSE
30	244.5	6.2	1877	1	PKC5_MOUSE
31	208	5.3	1680	1	FUR2_DROME
32	195.5	4.9	937	1	PAC4_RAT
33	184.5	4.7	913	1	PKC5_HUMAN

34	181	4.6	245	1	ERB2_MOUSE	P70424 mus musculus
35	180.5	4.5	1877	1	PKC5_RAT	P41413 rattus norv
36	176	4.5	634	1	ERBB_ALV	P00534 avian leuco
37	176	4.5	969	1	PAC4_HUMAN	P29122 homo sapien
38	174	4.4	604	1	ERBB_AVIER	P00535 avian eryth
39	171.5	4.3	2003	1	NTC4_HUMAN	O99466 homo sapien
40	168.5	4.3	3110	1	LMA2_HUMAN	P24043 homo sapien
41	164.5	4.2	3695	1	LMA2_HUMAN	O15230 homo sapien
42	164	4.1	540	1	ERBB_AVIED	P11273 avian eryth
43	162.5	4.1	1581	1	LMG3_MOUSE	O910b6 mus musculu
44	162	4.1	1173	1	TSP1_XENLA	P35448 xenopus lae
45	161	4.1	3084	1	LMA1_MOUSE	P19137 mus musculu

## ALIGNMENTS

RESULT 1  
ID ERB2\_HUMAN STANDARD: PRT: 1255 AA.  
AC P04626;  
DR 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)  
DE (p185erbB2) (NEU proto-oncogene) (C-erbB-2) (Tyrosine kinase-type cell  
DE surface receptor HER2) (MLN 19).  
GN ERBB2 OR HER2 OR NGL OR NEU.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Carnivora; Homiidae; Homo.  
OX NCBI\_Taxid:9606;  
RN [1]  
RP MEDLINE=86118663; PubMed=3003577;  
RA Yamamoto T., Ikawa S., Akiyama T., Semba K., Nomura N., Miyajima N.,  
RA Saito T., Toyoshima K.,  
RT "Similarity of protein encoded by the human c-erbB-2 gene to  
RT epidermal growth factor receptor.";  
RL Nature 319:230-234(1986).  
RN [2]  
RP MEDLINE=86070181; PubMed=2999974;  
RA Coussens L., Yang-Feng T.L., Liao Y.C., Chen E., Gray A.,  
RA McGrath J., Seeburg P.H., Libermann T.A., Schlessinger J.,  
RA Francke U., Levinson A., Ullrich A.,  
RT "Tyrosine kinase receptor with extensive homology to EGF receptor  
RT shares chromosomal location with neu oncogene.";  
RL Science 230:1132-1139(1985).  
RN [3]  
RP MEDLINE=737-1031 FROM N.A.  
RA MEDLINE=86016729; PubMed=2995967;  
RA Semba K., Kamata N., Toyoshima K., Yamamoto T.,  
RT "A v-erbB-related protooncogene, c-erbB-2, is distinct from the  
RT c-erbB-1/epidermal growth factor receptor gene and is amplified in a  
RT human salivary gland adenocarcinoma.";  
RL Proc. Natl. Acad. Sci. U.S.A. 82:6497-6501(1985).  
RN [4]  
RP VARIANTS VAL-654 AND VAL-655.  
RX MEDLINE=93194196; PubMed=8095488;  
RA Ehsani A., Low J., Wallace R.B., Wu A.M.,  
RT "Characterization of a new allele of the human ERBB2 gene by allele-  
RT specific competition hybridization.";  
RL Genomics 15:426-429(1993).  
RN [5]  
RP FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX,  
RP ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GP30 IS A  
RP POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-  
RP ALPHA AND AMPHIREGULIN.  
RN [6]  
RP CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
RP tyrosine phosphate.  
RN [7]  
RP SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS  
RP (POTENTIAL).  
RN [8]  
RP SUBCELLULAR LOCATION: Type I membrane protein.



CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE  
 CC RESIDUES (BY SIMILARITY).  
 CC -1- POLYMORPHISM: THERE ARE FOUR ALLELES DUE TO THE VARIATIONS IN  
 CC POSITIONS 654 AND 655. ALLELE B1 (654-ILE-ILE-655) HAS A FREQUENCY  
 CC OF 0.782; ALLELE B2 (654-ILE-VAL-655) HAS A FREQUENCY OF 0.206;  
 CC ALLELE B3 (654-VAL-VAL-655) HAS A FREQUENCY OF 0.012.  
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M11767; AAA35808.1; -;  
 DR EMBL: M11761; AAA35808.1; JOINED.  
 DR EMBL: M11762; AAA35808.1; JOINED.  
 DR EMBL: M11763; AAA35808.1; JOINED.  
 DR EMBL: M11764; AAA35808.1; JOINED.  
 DR EMBL: M11765; AAA35808.1; JOINED.  
 DR EMBL: M11766; AAA35808.1; JOINED.  
 DR EMBL: M11730; AAA75493.1; -;  
 DR EMBL: M12036; AAA35978.1; -;  
 DR EMBL: X03363; CAA27060.1; -;  
 DR PIR: A25491; A25491.  
 DR PIR: A24571; A24571.  
 DR HSSP: P11362; 1FGK.  
 DR Genew: HGNC:3430; ERBB2.  
 DR MIM: 164870; -;  
 DR InterPro: IPR000494; EGFR\_L\_domain.  
 DR InterPro: IPR000719; Euk\_kinase.  
 DR InterPro: IPR002174; Furin-like.  
 DR InterPro: IPR001245; Tyr\_kinase.  
 DR InterPro: IPR004019; YLP\_motif.  
 DR Pfam: PF00069; kinase; 1.  
 DR Pfam: PF00757; Furin-like; 1.  
 DR Pfam: PF01030; Recep\_L\_domain; 2.  
 DR Pfam: PF02757; YLP; 2.  
 DR Prodom: PD0000001; Euk\_kinase; 1.  
 DR SMART: SM00261; F0; 3.  
 DR SMART: SM00219; Tyrc; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 KM Transmembrane, Glycoprotein, Multigene family, Receptor, Signal;  
 KM transfease, Tyrosine-protein kinase; ATP-binding; Phosphorylation;  
 KM Polymorphism.  
 FT CHAIN 1 21 POTENTIAL.  
 FT 22 1255 RECEPTOR PROTEIN-TYROSINE KINASE ERBB-2.  
 FT DOMAIN 22 652 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 653 675 POTENTIAL.  
 FT DOMAIN 676 1255 CYTOPLASMIC (POTENTIAL).  
 FT 720 987 PROTEIN KINASE.  
 FT NP\_BIND 726 734 ATP (BY SIMILARITY).  
 FT BINDING 753 753 ATP (BY SIMILARITY).  
 FT ACT\_SITE 845 845 BY SIMILARITY.  
 FT DISULFID 195 204 BY SIMILARITY.  
 FT 199 212 BY SIMILARITY.  
 FT DISULFID 220 227 BY SIMILARITY.  
 FT 224 235 BY SIMILARITY.  
 FT DISULFID 236 244 BY SIMILARITY.  
 FT DISULFID 240 252 BY SIMILARITY.  
 FT 255 264 BY SIMILARITY.  
 FT DISULFID 268 295 BY SIMILARITY.  
 FT 299 311 BY SIMILARITY.  
 FT DISULFID 315 331 BY SIMILARITY.  
 FT 338 338 BY SIMILARITY.  
 FT DISULFID 511 520 BY SIMILARITY.  
 FT 515 528 BY SIMILARITY.  
 FT 531 540 BY SIMILARITY.  
 FT DISULFID 544 560 BY SIMILARITY.

FT DISULFID 563 576 BY SIMILARITY.  
 FT DISULFID 567 584 BY SIMILARITY.  
 FT DISULFID 587 596 BY SIMILARITY.  
 FT DISULFID 600 623 BY SIMILARITY.  
 FT DISULFID 626 634 BY SIMILARITY.  
 FT DISULFID 630 642 BY SIMILARITY.  
 FT MOD\_RES 1139 1139 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT MOD\_RES 1248 1248 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 124 124 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 187 187 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 571 571 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 629 629 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARIANT 654 654 I -> V.  
 FT VARIANT /FTID-VAR\_004077.  
 FT VARIANT I -> V.  
 FT VARIANT /FTID-VAR\_004078.  
 FT CONFLICT 1170 1170 P -> A (IN REF. 2).  
 FT SEQUENCE 1255 AA; 137909 MM; 39E9DFDA04DCF967 CRC64;  
 Query Match 95.5%; Score 3776; DB 1; Length 1255;  
 Best Local Similarity 67.9%; Pred. No. 6.9e-280; Mismatches 0; Indels 336; Gaps 1;  
 Matches 712; Conservative 0; Mismatches 0; Indels 336; Gaps 1;  
 QY 1 MELALACRMGLLALLPPGAASITQVCTGTDMLRLPASPETHLDMRLHLYOGCQVVGNTL 60  
 DB 1 MELALACRMGLLALLPPGAASITQVCTGTDMLRLPASPETHLDMRLHLYOGCQVVGNTL 60  
 QY 61 ELTYLPINASTLSFLODIOEVGYVLIHNOVROYLPRLRIVRGTOLEFEDNYALAVLDNG 120  
 DB 61 ELTYLPINASTLSFLODIOEVGYVLIHNOVROYLPRLRIVRGTOLEFEDNYALAVLDNG 120  
 QY 121 DPLNNTPTVPGASPGRLQLRSTELLKGVLIQNNPQLCYDTILMDIFKNNOLA 180  
 DB 121 DPLNNTPTVPGASPGRLQLRSTELLKGVLIQNNPQLCYDTILMDIFKNNOLA 180  
 QY 181 LTLIDTNRSRACHPCSPMGKSRGESSSEDCQSLFTVYAGGARGKGLPTDCHEQC 240  
 DB 181 LTLIDTNRSRACHPCSPMGKSRGESSSEDCQSLFTVYAGGARGKGLPTDCHEQC 240  
 QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDFESMPNREGRYTFGASCVTACP 300  
 DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDFESMPNREGRYTFGASCVTACP 300  
 QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDFESMPNREGRYTFGASCVTACP 300  
 DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDFESMPNREGRYTFGASCVTACP 300  
 QY 301 YNLTSTVGSCTLVCPRLHNOEYVAEDGTORCEKSKCAVCYGLGEMHLREVRATVSAN 360  
 DB 301 YNLTSTVGSCTLVCPRLHNOEYVAEDGTORCEKSKCAVCYGLGEMHLREVRATVSAN 360  
 QY 361 IOEPAGCKITFGSLAFIPESFDGDPASNTAPLQPEQLOVETLEITGYLYISAMPDLP 420  
 DB 361 IOEPAGCKITFGSLAFIPESFDGDPASNTAPLQPEQLOVETLEITGYLYISAMPDLP 420  
 QY 421 DLSYFQMLQYIRGRILHNGAVSLTGLGISWLGSLRLSRLSGLALIHNTHLCEVHTV 480  
 DB 421 DLSYFQMLQYIRGRILHNGAVSLTGLGISWLGSLRLSRLSGLALIHNTHLCEVHTV 480  
 QY 481 PMQDLFNPQALLHTNRREDECVGGLACHQLCAGHCHWGPRQVCVNCOSQFLRQEC 540  
 DB 481 PMQDLFNPQALLHTNRREDECVGGLACHQLCAGHCHWGPRQVCVNCOSQFLRQEC 540  
 QY 541 VEECRVIOGLPREYVNAHCLPCHPEQOPONGSVTCGPADQVACAHKDDPPFCVARG 600  
 DB 541 VEECRVIOGLPREYVNAHCLPCHPEQOPONGSVTCGPADQVACAHKDDPPFCVARG 600  
 QY 601 PSQVKKPDLSTYMPKPFDEGACQPCINCTHSCVDLDDKGCAPAEQASPLTSTISAVG 660  
 DB 601 PSQVKKPDLSTYMPKPFDEGACQPCINCTHSCVDLDDKGCAPAEQASPLTSTISAVG 660  
 QY 654 ----- 653  
 DB 661 ILVVVLGVVFGILIKRRQOKIRKTYTMRLLQETELVEPLTPSGAMPNOAMRILKTEL 720







```

FT CHAIN 22 1254 RECEPTOR PROTEIN-TYROSINE KINASE ERB-B-2.
FT DOMAIN 22 652 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 653 675 POTENTIAL.
FT DOMAIN 676 1254 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 158 368 CYS-RICH.
FT DOMAIN 472 644 CYS-RICH.
FT DOMAIN 720 987 PROTEIN KINASE.
FT NP_BIND 726 734 ATP (BY SIMILARITY).
FT BINDING 753 753 ATP (BY SIMILARITY).
FT ACT_SITE 845 845 BY SIMILARITY.
FT DISULFID 195 204 BY SIMILARITY.
FT DISULFID 199 212 BY SIMILARITY.
FT DISULFID 236 244 BY SIMILARITY.
FT DISULFID 240 252 BY SIMILARITY.
FT DISULFID 255 264 BY SIMILARITY.
FT DISULFID 268 295 BY SIMILARITY.
FT DISULFID 299 311 BY SIMILARITY.
FT DISULFID 315 331 BY SIMILARITY.
FT DISULFID 334 338 BY SIMILARITY.
FT DISULFID 511 520 BY SIMILARITY.
FT DISULFID 515 528 BY SIMILARITY.
FT DISULFID 531 540 BY SIMILARITY.
FT DISULFID 544 560 BY SIMILARITY.
FT DISULFID 563 576 BY SIMILARITY.
FT DISULFID 567 584 BY SIMILARITY.
FT DISULFID 587 596 BY SIMILARITY.
FT DISULFID 600 623 BY SIMILARITY.
FT DISULFID 626 634 BY SIMILARITY.
FT MOD_RES 1139 1139 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 1247 1247 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 187 187 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 530 530 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 571 571 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 629 629 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 658 658 V -> E (IN ONCOGENIC NEU).
FT VARIANT 659 659 V -> E (IN ONCOGENIC NEU).
SQ SEQUENCE 1254 AA; 138252 MW; 974C3791C21F2B81 CRC64;

Query Match 80.6%; Score 3185; DB 1; Length 1254;
Best Local Similarity 57.5%; Pred. No. 7.5e-235;
Matches 603; Conservative 36; Mismatches 73; Indels 336; Gaps 1;

```

```

QY 421 DLSTFQNLQVIRGRITLHNGAVSLTLOGISWLGSLRLREIGSGALITHNHILCFVHTV 480
DB 421 DLSTFQNLQVIRGRITLHNGAVSLTLOGISWLGSLRLREIGSGALITHNHILCFVHTV 480
QY 481 PMDQLFRPHQALLHTANRPDECEGEGEGLACHQGLARHCWGPSTOCVNCSPFRGQEC 540
DB 481 PMDQLFRPHQALLHTANRPDECEGEGEGLACHQGLARHCWGPSTOCVNCSPFRGQEC 540
QY 541 VEECVNLQGLPREVYVNAHRCLEPCHECOPQNGSVTCFGEADQCVACAHYKDPPECVARC 600
DB 541 VEECVNLQGLPREVYVNAHRCLEPCHECOPQNGSVTCFGEADQCVACAHYKDPPECVARC 600
QY 541 VEECVNLQGLPREVYVNAHRCLEPCHECOPQNGSVTCFGEADQCVACAHYKDPPECVARC 600
DB 541 VEECVNLQGLPREVYVNAHRCLEPCHECOPQNGSVTCFGEADQCVACAHYKDPPECVARC 600
QY 601 PSQVAPDLSTYPMIMKFPPEEGACQPCPTNCCHSCVDLDDKGRAPQRRASPLTS 653
DB 601 PSQVAPDLSTYPMIMKFPPEEGACQPCPTNCCHSCVDLDDKGRAPQRRASPLTS 653
QY 654 ----- 653
DB 654 ----- 653
QY 654 ----- 653
DB 654 ----- 653
QY 721 RKVKVLGSGAGTGYKGIWIPDGENVKIPVAIKVLRNTSPRANKELIDEAYVAGLSP 780
DB 721 RKVKVLGSGAGTGYKGIWIPDGENVKIPVAIKVLRNTSPRANKELIDEAYVAGLSP 780
QY 654 ----- 653
DB 654 ----- 653
QY 781 YVSRLLGICLSTVOLVTQMLPFGCLLDHVRHGRGLSGQDLNMCVOIAKMSYLEDVR 840
DB 781 YVSRLLGICLSTVOLVTQMLPFGCLLDHVRHGRGLSGQDLNMCVOIAKMSYLEDVR 840
QY 654 ----- 653
DB 654 ----- 653
QY 841 LVHRLAARLVLYKSPNHYKITDFGLARLIDETEFHADGKVPYKMALESILRRFT 900
DB 841 LVHRLAARLVLYKSPNHYKITDFGLARLIDETEFHADGKVPYKMALESILRRFT 900
QY 654 ----- 653
DB 654 ----- 653
QY 901 HOSDVSYGVYVWELMTFGAKPYDGPAREIPDLLEKGBRLPQPICTIDYVIMWKCMM 960
DB 901 HOSDVSYGVYVWELMTFGAKPYDGPAREIPDLLEKGBRLPQPICTIDYVIMWKCMM 960
QY 654 ----- 653
DB 654 ----- 653
QY 961 IDSECRPFRELSEFSRMAPDPQRFVYIOMEDLSPSSPLDSTFYSRLLEDMDGLDVA 1020
DB 961 IDSECRPFRELSEFSRMAPDPQRFVYIOMEDLSPSSPLDSTFYSRLLEDMDGLDVA 1020
QY 685 EEYLVPOQGFPCDPAPGAGGVHHRHR 712
DB 685 EEYLVPOQGFPCDPAPGAGGVHHRHR 712
QY 1021 EEYLVPOQGFPCDPAPGAGGVHHRHR 1048
DB 1021 EEYLVPOQGFPCDPAPGAGGVHHRHR 1048

RESULT 4
EGFR_MOUSE STANDARD: PRT: 1210 AA.
AC 001279;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Epidermal growth factor receptor precursor (EC 2.7.1.112).
GN EGFR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Balb/c; TISSUE=Liver;
RX MEDLINE=93026370; PubMed=1408137;
RA Avivi A., Skorecki K., Yaron A., Givol D.;
RT "Promoter region of the murine fibroblast growth factor receptor 2
RT (bek/KGFR) gene."
RT Oncogene 7:1957-1962(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Balb/c, and CD-1; TISSUE=Liver, and Decidua;
RX MEDLINE=93126380; PubMed=7678348;
RA Paria B.C., Das S.K., Andrews G.K., Dey S.K.;
RT "Expression of the epidermal growth factor receptor gene is regulated

```



QY 244 CTGPKHSDCLACLNHNSHSGICELHCPALVYNTDFFESMPNPEGRYTGACVCTACPYNY 303  
 DB 240 CTGPRESDCLVQKRFQDBATCKDKCPRPLMLYNPTTYQMDVNBEGKYSRGATCVKCKCPNXY 299  
 QY 304 LSTDVSGCTLVCPPLHNOVYTAEDTQRCCKSKPCARVCYGLGMEHLREAVRASANTIOE 363  
 DB 300 VYTDHGSCVRCAGPDYEVY-EDDGIRCKCKDGCRCRKYCNNGIGIEFDDTSLINATNIKH 358  
 QY 364 FAGCKKIGSLAFIPESDGPASPTAPLOEQLOVFEETLEITGKYLXISAMPDLSPLDS 423  
 DB 359 FKYCAISGDHLPLVAVKGSFTPTPLDRELEITLVEITGELLIQAMPDMWTDLH 418  
 QY 424 VFONLQVIRGTHLNGAVSLTLOGISMLGSLRELSGSLALIHHTHLCFVTTVPMD 483  
 DB 419 AFENELIRGRTRKQHGOSFLAVGINTSLGRLSKETSDSDVITSGNRNLCTANTIMWK 478  
 QY 484 QLFNPNHQAALLHTANRPDECEVGEGLACHQCLCARGHCGPPTQCVNCSQPLRGCEVVE 543  
 DB 479 KLFETPNQKTKIMNNRAEKDCKAVNHVCNPLCSSGECGWPBPCDVCQNTSRGREGVEK 538  
 QY 544 CRVIGLQPREVVNARHCLPCHEPCOPONGSVTCGPREPDOVACAHYNDPPVCARCSG 603  
 DB 539 CNLIEGPREVEVSECTIQCHPECLPQAMNITCTGRGPDNCTQCAHYIDGHCYVKTCPAG 598  
 QY 604 VKPDLSTYPIKWPDEEGACQPCPINCTHSCVDLDDKGC 642  
 DB 599 IMGENNTL-VWKYADANNVCHLCHANCTYGCAGPGLQGC 636

## RESULT 5

EGFR\_HUMAN

ID EGFR\_HUMAN STANDARD: PRT: 1210 AA.

AC P00533; P06268; Q14225; Q9UMD7; Q9UMD8; Q9UMG5; Q92795; 000732;  
 AC 000688; Q9B852; Q9B2C9; Q9B2X1; Q9H3C9;  
 DT 21-JUL-1986 (Rel. 01, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Epidermal growth factor receptor precursor (EC 2.7.1.112) (Receptor  
 protein-tyrosine kinase EtdB-1).

GN EGFR OR EKBBI.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI\_TaxId=9606;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORM 1).

RX MEDLINE-84219729; PubMed-6328312;  
 RA Ullrich A., Coussens L., Hayflick J.S., Dull T.J., Gray A., Tam A.W.,  
 RA Lee J., Yarden Y., Libermann T.A., Schlessinger J., Downward J.,  
 RA Mayes E.L.V., Whittle N., Waterfield M.D., Seeburg P.H.;RT "Human epidermal growth factor receptor cDNA sequence and aberrant  
 RT expression of the amplified gene in A431 epidermoid carcinoma cells."

RL Nature 309:418-425(1984).

RN [2]

RP SEQUENCE FROM N.A. (ISOFORM 2).

RX MEDLINE-95382957; PubMed-7654368;  
 RA Ilekis J.V., Stark B.C., Scoccia B.;RT "Possible role of variant RNA transcripts in the regulation of  
 RT epidermal growth factor receptor expression in human placenta.";

RL Mol. Reprod. Dev. 41:149-156(1995).

RN [3]

RP SEQUENCE FROM N.A. (ISOFORM 2).

RX MEDLINE-97078686; PubMed-8918811;  
 RA Reller J.L., Maible N.J.;RT "A 1.8 kb alternative transcript from the human epidermal growth  
 RT factor receptor gene encodes a truncated form of the receptor.";

RL Nucleic Acids Res. 24:4050-4056(1996).

RN [4]

RP SEQUENCE FROM N.A. (ISOFORM 2).

RX MEDLINE-97256547; PubMed-9103388;

RA Ilekis J.V., Gariti J., Niederberger C., Scoccia B.;

RT "Expression of a truncated epidermal growth factor receptor-like  
 RT protein (TEGFR) in ovarian cancer."

RL Gynecol. Oncol. 65:36-41(1997).

RN [5]

RP SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).

RC TISSUE-Placenta;

RX MEDLINE-21100872; PubMed-11161793;  
 RA Reller J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielsen A.J.,  
 RA Scheel Sinclair C., Pearshall R.S., Green P.J., Yee D., Lampland A.L.,  
 RA Balasubramaniam S., Crossley T.D., Magnuson T.R., James C.D.,  
 RA Maible N.J.;

RT "Comparative genomic sequence analysis and isolation of human and  
 RT mouse alternative EGFR transcripts encoding truncated receptor  
 RT isoforms.";

RL Genomics 71:1-20(2001).

RN [6]

RP SEQUENCE OF 575-687 FROM N.A.

RA Reller J.L., Threadgill D.W., Danielsen A.J., Scheel C.M.,  
 RA Lampland A.L., Balasubramaniam S., Crossley T.O., Magnuson T.R.,  
 RA Maible N.J.;

RT "Human and mouse alternative EGFR transcripts encoding only the  
 RT extracellular domain of the receptor.";

RL Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.

RN [7]

RP SEQUENCE OF 713-924 FROM N.A.

RX MEDLINE-84196372; PubMed-6326261;  
 RA Lin C.R., Chen W.S., Kruliger W., Stolarsky L.S., Weber W.,  
 RA Evans R.M., Verma I.M., Gill G.N., Rosenfield M.G.;

RT "Expression cloning of human EGF receptor complementary DNA: gene  
 RT amplification and three related messenger RNA products in A431  
 RT cells."

RL Science 224:843-848(1984).

RN [8]

RP SEQUENCE OF 150-962 FROM N.A.

RX MEDLINE-84245835; PubMed-6330563;  
 RA Xu Y.H., Ishii S., Clark A.J.L., Sullivan M., Wilson R.K., Ma D.P.,  
 RA Roe B.A., Merlino G.T., Pastan I.;

RT "Human epidermal growth factor receptor cDNA is homologous to a  
 RT variety of RNAs overproduced in A431 carcinoma cells."

RL Nature 309:806-810(1984).

RN [9]

RP SEQUENCE OF 1028-1210 FROM N.A.

RX MEDLINE-85046483; PubMed-6093780;  
 RA Simmen F.A., Gope M.L., Schultz T.Z., Wright D.A., Carpenter G.,  
 RA O'Malley B.W.;

RT "Isolation of an evolutionarily conserved epidermal growth factor  
 RT receptor cDNA from human A431 carcinoma cells."

RL Biochem. Biophys. Res. Commun. 124:125-132(1984).

RN [10]

RP SEQUENCE OF 1-29 FROM N.A.

RX MEDLINE-88217333; PubMed-3329716;  
 RA Haley J.D., Whittle N., Bennett P., Kinchington D., Ullrich A.,  
 RA Waterfield M.D.;

RT "The human EGF receptor gene: structure of the 110 kb locus and  
 RT identification of sequences regulating its transcription."

RL Oncogene Res. 1:373-396(1987).

RN [11]

RP SEQUENCE OF 1-29 FROM N.A.

RX MEDLINE-91107677; PubMed-1988448;  
 RA Haley J.D., Waterfield M.D.;

RT "Contributory effects of de novo transcription and premature  
 RT transcript termination in the regulation of human epidermal growth  
 RT factor receptor proto-oncogene RNA synthesis."

RL J. Biol. Chem. 266:1746-1753(1991).

RN [12]

RP SEQUENCE OF 1-29 FROM N.A.

RX MEDLINE-85270438; PubMed-2991899;  
 RA Ishii S., Xu Y.H., Stratton R.H., Roe B.A., Merlino G.T., Pastan I.;

RT "Characterization and sequence of the promoter region of the human  
 RT epidermal growth factor receptor gene."

RL Proc. Natl. Acad. Sci. U.S.A. 82:4920-4924(1985).

RN [13]

SEQUENCE OF 540.  
Submitted (SEP-1997) to the SWISS-PROT data bank.  
[14]  
RECEPTOR ACTIVITY.  
MEDLINE=84191554; PubMed=6325948;  
Mroczkowski B., Mosig G., Cohen S.;  
"App-stimulated interaction between epidermal growth factor receptor  
and supercoiled DNA.";  
Nature 309:270-273(1984).  
[15]  
PHOSPHORYLATION.  
MEDLINE=89278137; PubMed=2543678;  
Margolis B.L., Lax I., Kris R., Dombalagian M., Honegger A.M.,  
Hovik R., Givol D., Ullrich A., Schlessinger J.;  
"All autophosphorylation sites of epidermal growth factor (EGF)  
receptor and HER2/neu are located in their carboxyl-terminal tails.  
Identification of a novel site in EGF receptor.";  
J. Biol. Chem. 264:10667-10671(1989).  
[16]  
CARBOHYDRATE-LINKAGE SITES ASN-128; ASN-175; ASN-413; ASN-444 AND  
ASN-528.  
MEDLINE=96398132; PubMed=8962717;  
Smith K.D., Davies M.J., Bailey D., Renouf D.V., Hounsell E.F.;  
"Analysis of the glycosylation patterns of the extracellular domain of  
the epidermal growth factor receptor expressed in Chinese hamster  
ovary fibroblasts.";  
Growth Factors 13:121-132(1996).  
[17]  
CARBOHYDRATE-LINKAGE SITES ASN-56; ASN-352; ASN-361; ASN-568 AND  
ASN-603.  
MEDLINE=20198209; PubMed=10731668;  
Saito C., Kim Y.-H., Abe Y., Saito K., Yokoyama S., Kohda D.;  
"Characterization of the N-oligosaccharides attached to the atypical  
Asn-X-Cys sequence of recombinant human epidermal growth factor  
receptor";  
J. Biochem. 127:65-72(2000).  
[18]  
PARTIAL SEQUENCE, AND DISULFIDE BONDS.  
MEDLINE=98225196; PubMed=9556602;  
Abe Y., Oaka M., Inagaki F., Iax I., Schlessinger J., Kohda D.;  
"Disulfide bond structure of human epidermal growth factor receptor.";  
J. Biol. Chem. 273:11150-11157(1998).  
[19]  
REVIEW.  
MEDLINE=87297456; PubMed=3039909;  
Carpenter G.;  
"Receptors for epidermal growth factor and other polypeptide  
mitogens.";  
Annu. Rev. Biochem. 56:881-914(1987).  
-I- FUNCTION: Receptor for EGF, but also for other members of the EGF  
family, as TGF-alpha, amphiregulin, betacellulin, heparin-binding  
EGF-like growth factor, GP30 and vaccinia virus growth factor. Is  
involved in the control of cell growth and differentiation.  
-I- FUNCTION: Isoform 2/truncated isoform may act as an antagonist.  
-I- CATALYTIC ACTIVITY: ATP + a protein Tyrosine = ADP + protein  
tyrosine phosphate.  
-I- SUBCELLULAR LOCATION: Type I membrane protein. Isoform 2 is  
secreted.  
-I- ALTERNATIVE PRODUCTS: 4 isoforms; 1/p170 (shown here), 2/p60/  
truncated isoform/TEGFR, 3/p110 and 4; are produced by  
alternative splicing.  
-I- TISSUE SPECIFICITY: Expressed in placenta. Isoform 2 is also  
expressed in ovarian cancers.  
-I- MISCELLANEOUS: Binding of EGF to the receptor leads to  
dimerization, internalization of the EGF-receptor complex,  
induction of the tyrosine kinase activity, stimulation of cell DNA  
synthesis, and cell proliferation.  
-I- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/cc> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

Query_Match	38.7%	Score 1532;	DB 1;	Length 1210;
Best Local Similarity	42.9%	Pred. No. 6.6e-109;		
Matches	283;	Conservative 108;	Mismatches 241;	Indels 28; Gaps 6;
Dr EMBL; X00588; CA25240.1; -				
Dr EMBL; U95089; AAB53063.1; -				
Dr EMBL; U48722; AAC50802.1; -				
Dr EMBL; U48723; AAC50804.1; -				
Dr EMBL; U48724; AAC50796.1; -				
Dr EMBL; U48725; AAC50797.1; -				
Dr EMBL; U48726; AAC50798.1; -				
Oy 11 L1ALALPPGAA--SYOVCTGTDMDKLRLPASPETHIDMLRHLRYHIOGCOVOGNILETYLPIN 68				
Db 14 L1AALCPASRALEEKVKVOGTSNKLTQGTGEDPHDLISLQRMENNEVYLGNIETIYVQRN 73				
Oy 69 ASLSFLQDIOEVGYVLIAHNOVRCVPLQRLRYVGTOLFEDNVALAVLNDGDPINNTTP 128				
Db 74 YDLSEFLKTIQEVAGYVLIALLNTVTERIIPLENQIILGNMYEYSVALAVLSND----- 126				
Oy 129 VTGASPGGLRLQLSLTEILKGVLIQRNLQCYODTIIMKDIIFHKNNQALATLIDTNR 188				
Db 127 --AAKTKIETKELPMNLDELHGAVFSSNPMLCVESI-RQMDIVSYDPSLSMNSMDFQNH 183				
Oy 189 SRACPPCSPMCKSGKSGMESSDCCSLRTYVAGGCA-RCKGPLRTDCCHEGCAAGCTGP 247				
Db 184 LGSCKCPDPCPCNGSCWAGENECKKLKIIICAOCSGRCKRSPSDCHNCAGCGCP 243				
Oy 248 KHSDCIACLFHNHSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVPYNYLSTD 307				
Db 244 RESDCLVCRKFEDEATCKDTCPLMLNPTTYQMDVNPBGKYSFGATGCVKCKPRNYVTD 303				
Oy 308 VGSCLVCPRLNHQETADGTQRCRCKSKPCARVCYGLGMELRLVRAVTSANIODEFGC 367				
Db 304 HGSCVRACGADSYEM-EDGVAKCKCKGCPCKVYNGIGIEGKDSLIMATNIKHFKN 362				
Oy 368 KKFISFLAFPEFSFGDPASNTAPRLQPPOLQVFEETLEITGLYIYSAMPDLSPLDSVFQN 427				
Db 363 TSISSDHLHLPYAFGDSFTHTPRLDPOELDLTKYKELTGLLQLQAMPENTDIAHEN 422				
Oy 428 LQVIRGRILHNGAVSLTLOGIGISWLGRLSLRELSSGLALHNHNLHCFVHTVPDQLFR 487				
Db 423 LEIRGRKROHQFSLAVVSLNITSLGRSLKEISDGVITISGNKILCAVNTINKKLEF 482				
Oy 488 NPHQALLHTANRPEDECYGEGLACIQLCARHGCMRGPRLQCYNCQFLRGDCYVEECRL 547				
Db 483 TSGQTKTIIISNGENSCATGVCHALOSPESCMGPEPDDCVSCSNVSGRCVCKML 542				
Oy 548 OGLPEEVYANARCLCHPEECOPQNSVYVCEGEADQCVACVAKYKPPFCVAPCPSGVAPD 607				
Db 543 DEEPPEFVENSECICHCBLDPOAMNITITCTGPRNCIQCANI IDGPHCVKTCRPGVNGE 602				
Oy 608 LSYMPITFFPDEEGACOPPCINCTHSCVDLDKCGCPAEQRASPLTSQNDLGPASPLDST 667				
Db 603 NNTL-VMKYADAGHCHLCHPCTGTGCGTGGLEGGPTN-----GPKIPSLAT 648				

RESULT 6  
EGFR\_CHICK  
ID EGFR\_CHICK STANDARD: PRT: 703 AA.  
AC P13387;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Epidermal growth factor receptor precursor (EC 2.7.1.112) (CER)  
DE (Fragment).  
GN EGFR.  
OS Gallus gallus (Chicken).

CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
CC	Gallus.
OX	NCBI_TaxID=9031;
RM	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=88261272; PubMed=3260329;
RA	Lax I., Johnson A., Hawk R., Sap J., Bellot F., Winkler M.,
RA	Ullrich A., Vennstrom B., Schlessinger J., Glyvol D.;
RT	"Chicken epidermal growth factor (EGF) receptor: cDNA cloning,
RT	expression in mouse cells, and differential binding of EGF and
RT	transforming growth factor alpha."
RL	Mol. Cell Biol. 8:1970-1978(1988).
CC	-1- FUNCTION: THE EGF RECEPTOR MEDIATES THE BIOLOGICAL SIGNAL OF EGF,
CC	AND ALSO OF TGF-ALPHA, AMPHIREGULIN, HEPARIN-BINDING EGF, GP30 AND
CC	VACCINIA VIRUS GROWTH FACTOR (BY SIMILARITY).
CC	-1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC	tyrosine phosphate.
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.
CC	-1- MISCELLANEOUS: Binding of EGF to the receptor leads to
CC	dimerization, internalization of the EGF-receptor complex,
CC	induction of the tyrosine kinase activity, stimulation of cell DNA
CC	synthesis, and cell proliferation.
CC	-1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
CC	-----
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch">http://www.isb-sib.ch/announce/</a>
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
CC	-----
DR	EMBL: M20386; AAA48760.1; .
DR	InterPro: IPR000494; EGFR_L_domain.
DR	InterPro: IPR000719; Euk-Pkinase.
DR	InterPro: IPR002174; Furin-like.
DR	InterPro: IPR001245; Tyr_Pkinase.
DR	Pfam: PF00757; Furin-like; 1.
DR	Pfam: PF01030; Recep_L_domain; 2.
DR	SMART: SMO0261; Fu_4
DR	PROSITE: PS00107; PROTEIN_KINASE_ATP_PARTIAL.
DR	PROSITE: PS00109; PROTEIN_KINASE_TYR_PARTIAL.
DR	PROSITE: PS50011; PROTEIN_KINASE_DOM_PARTIAL.
KW	Transmembrane; Glycoprotein; Receptor; Signal; Transferase;
KW	Tyrosine-protein kinase; ATP-binding; Phosphorylation.
FT	SIGNAL
FT	1
FT	30
FT	CHAIN
FT	31
FT	>703
FT	DOMAIN
FT	31
FT	654
FT	TRANSMEM
FT	655
FT	667
FT	POTENTIAL.
FT	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN
FT	668
FT	>703
FT	DISULFID
FT	197
FT	206
FT	201
FT	214
FT	DISULFID
FT	222
FT	230
FT	DISULFID
FT	226
FT	238
FT	DISULFID
FT	239
FT	247
FT	DISULFID
FT	243
FT	255
FT	DISULFID
FT	258
FT	267
FT	DISULFID
FT	271
FT	298
FT	DISULFID
FT	302
FT	314
FT	DISULFID
FT	318
FT	333
FT	DISULFID
FT	336
FT	340
FT	DISULFID
FT	513
FT	522
FT	DISULFID
FT	517
FT	530
FT	DISULFID
FT	533
FT	542
FT	DISULFID
FT	546
FT	562
FT	DISULFID
FT	565
FT	581
FT	DISULFID
FT	569
FT	589
FT	DISULFID
FT	592
FT	601
FT	DISULFID
FT	605
FT	627
FT	DISULFID
FT	630
FT	638
FT	DISULFID
FT	634
FT	646
FT	CARBOHYD
FT	134
FT	134
FT	EPIDERMAL GROWTH FACTOR RECEPTOR.
FT	EXTRACELLULAR (POTENTIAL).
FT	POTENTIAL.
FT	CYTOPLASMIC (POTENTIAL).
FT	DISULFID
FT	197
FT	206
FT	201
FT	214
FT	DISULFID
FT	222
FT	230
FT	DISULFID
FT	226
FT	238
FT	DISULFID
FT	239
FT	247
FT	DISULFID
FT	243
FT	255
FT	DISULFID
FT	258
FT	267
FT	DISULFID
FT	271
FT	298
FT	DISULFID
FT	302
FT	314
FT	DISULFID
FT	318
FT	333
FT	DISULFID
FT	336
FT	340
FT	DISULFID
FT	513
FT	522
FT	DISULFID
FT	517
FT	530
FT	DISULFID
FT	533
FT	542
FT	DISULFID
FT	546
FT	562
FT	DISULFID
FT	565
FT	581
FT	DISULFID
FT	569
FT	589

Query Match	Best Local Similarity	Score	DB 1;	Length	703;
Matches	286;	Conservative	103;	Mismatches	235;
				Indels	23;
				Gaps	
FT CAROHD 190	190	N-LINKED (GLCNAc . . .)	(POTENTIAL)		
FT CAROHD 200	200	N-LINKED (GLCNAc . . .)	(POTENTIAL)		
FT CAROHD 359	359	N-LINKED (GLCNAc . . .)	(POTENTIAL)		
FT CAROHD 368	368	N-LINKED (GLCNAc . . .)	(POTENTIAL)		
FT CAROHD 420	420	N-LINKED (GLCNAc . . .)	(POTENTIAL)		
FT CAROHD 573	573	N-LINKED (GLCNAc . . .)	(POTENTIAL)		
FT CAROHD 578	578	N-LINKED (GLCNAc . . .)	(POTENTIAL)		
FT CAROHD 613	613	N-LINKED (GLCNAc . . .)	(POTENTIAL)		
FT CAROHD 633	633	N-LINKED (GLCNAc . . .)	(POTENTIAL)		
FT CAROHD 648	648	N-LINKED (GLCNAc . . .)	(POTENTIAL)		
FT NON_TER 703	703	N-LINKED (GLCNAc . . .)	(POTENTIAL)		
SQ SEQUENCE 703 AA:	703	AF22DE1B735A690	CRC64:		
Query Match	38.2%;	Score	1509.5;	DB 1;	Length 703;
Best Local Similarity	44.2%;	Pred. No.	1.8e-107;		
Matches	286;	Conservative	103;	Mismatches	235;
				Indels	23;
				Gaps	
8 RRGLLALLPPAA-----STVCTGIDMKRLPASPTHLMDLRLHYGCGVYQGNIE 61					
13 RGAAYLVLLLEVALCSAVERKKVCGGTNNKLTOLGHNEDHTSLDRMYNCEVLSNIE 72					
62 LUYLPNALSFLDIOEVOGVLLAHNOVROPLORLRTVGTOLFEENYALAVLND 121					
73 IIVVENRDPLFTLQIEAGVTLALNAVYIPLENLDITGNVLYLDSFLAVALSYNH 132					
122 PLNNTPTVYGASPGGLRELTLRLTEILKGVLIORNPOLCYODPTLMKDIFKKNQNAL 181					
133 -NNKQ-----GIRELPMKRSLSEILNGVKISNNPRLCMMDYIWMDDIDTSRK -PL 182					
182 TLID-TNRSRACHPCSPMCKSGSRGSESSDCOSLFTFYCAGCA-RCKGPLPTDCCHEQ 239					
183 IVLIDPASNLSLSCPKCHPNCETEDHCAGARONQVTLVLAOCQSGRCGKVPSPDCINQ 242					
240 CAAGCGPKHSPCLACIHNHSGICELHCPALVYTNTPDEEPMRDEGRTYRGASCVNAC 299					
243 CAAGCTGPRSDCLACRFRDADTKDTPRLVLTNPPTYOMDVNDEGYSFGATCVBRC 302					
300 PNYLSTVDGSCITLCPRLNDEVTAEDETQRCESCKSPCARVCYGLGEMHLEVRNAVISA 359					
303 PNNYVTVHDGSCVRCNDTTEYV-BENGVRRKKKKDGLCSKYCNIGIGELKGIISNANT 361					
360 NIOEPAGCKKIFGSLAFPESEFDGRPASNTAPLAPBOULOVFETLEBTGYLYISAMPDSL 419					
362 NIDSEKCKTKKINGDVISLPVAFGLGAFYKTLPLDPRKLDVFEVTEYISGFLLIQAMPNA 421					
420 PLSVFNOLVIRGVLNHNAGVSLTLOGIGISWLGRLSRLREGSGALILHNHTHLCFVHT 479					
422 TDLVFEENLETLRGKTKQGYSLVNVNKLISLGRSLKESISDGDIALMKKNLCYADT 481					
480 VPMDDLFRRNPOLALHTANRPDECEVSGEGLACHOLCARHCKWSPRTQCVNCSQFLRGE 539					
482 MMVRSLEFNTQSKTKIQRNNKNDCTADRHVCDPLCSDVCGCGPGRPHCFGRFESRQKE 541					
540 CVECEKRVLOGLRFREYVNAHRLCPRECOFONG---SYTCSEPRADQVACAHYADRPFC 596					
542 CYKQCNILQGEERERDSKCLPCHSECLVONSTAYNTTSCSPGRGDHCKAHFIDGPHC 601					
597 VARGSGVYKPRDLSYMPIMKFRPEDEBACOPRCINCHSCVDLDDKCP 643					
602 VACAPAGVLGENDTL-VMKYADANAVCOLCHPNCYTRCGKGPGLBECR 647					
RESULT 7					
ERR4_HUMAN					
ID ERR4_HUMAN	STANDARD:	PRF:	1308 AA.		
AC Q15303:					
DT 15-DEC-1998 (Rel. 37, Created)					
DT 15-DEC-1998 (Rel. 37, Last sequence update)					
DT 15-JUN-2002 (Rel. 41, Last annotation update)					
DE Receptor protein-tyrosine kinase erbB-4 precursor (EC 2.7.1.112)					
DE (p180erbB) (Tyrosine kinase-type cell surface receptor HER4).					
OS Homo sapiens (Human).					



OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM JM-A).  
 RC TISSUE=breast carcinoma;  
 RA MEDLINE=93189574; PubMed=8383326;  
 RA Plozman G.D., Culouscou J.-M., Whitney G.S., Green J.M., Carlton G.W.,  
 RA Foy L., Neuberger M.G., Shoyab M.;  
 RA "ligand-specific activation of HER4/p180erbB4, a fourth member of the  
 RT epidermal growth factor receptor family";  
 RT Proc. Natl. Acad. Sci. U.S.A. 90:1746-1750(1993).  
 RN (2)  
 RP SEQUENCE FROM N.A. (ISOFORMS JM-A AND JM-B).  
 RC TISSUE=Fetal brain;  
 RX MEDLINE=97476287; PubMed=9334263;  
 RA Elenius K., Corfas G., Paul S., Choi C.J., Rio C., Plozman G.D.,  
 RA Klagsbrun M.;  
 RA "A novel juxtamembrane domain isoform of HER4/erbB4. Isoform-specific  
 RT tissue distribution and differential processing in response to  
 RT phorbol ester";  
 RT J. Biol. Chem. 272:26761-26768(1997).  
 RL J. Biol. Chem. 272:26761-26768(1997).  
 CC -1- FUNCTION: SPECIFICALLY BINDS AND IS ACTIVATED BY NEUREGULINS, NRG-  
 CC 2, NRG-3, HEPARIN-BINDING EGF-LIKE GROWTH FACTOR, BETAHELULIN AND  
 CC NTAR. INTERACTION WITH THESE FACTORS INDICES CELL DIFFERENTIATION.  
 CC NOT ACTIVATED BY EGF, TGF-A, AND AMPHIREGULIN.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -1- SUBUNIT: HOMODIMER OR HETERODIMER WITH EACH OF THE OTHER ERBB  
 CC RECEPTORS (POTENTIAL).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: JM-A (SHOWN HERE) AND JM-B;  
 CC ARE PRODUCED BY ALTERNATIVE SPLICING. THE TWO ISOFORMS DIFFER  
 CC FUNCTIONALLY IN THEIR RESPONSE TO PHORBOL ESTER. JM-A IS PROCESSED  
 CC BUT NOT JM-B. SO THEY RESPECTIVELY REPRESENT CLEAVABLE AND  
 CC NONCLEAVABLE FORMS OF THE RECEPTOR. BOTH ISOFORMS ARE EXPRESSED IN  
 CC CEREBELLUM, BUT ONLY THE JM-B ISOFORM IS EXPRESSED IN THE HEART.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED AT HIGHEST LEVELS IN BRAIN, HEART,  
 CC KIDNEY, IN ADDITION TO SKELETAL MUSCLE, PARATHYROID, CEREBELLUM,  
 CC PITUITARY, SPLEEN, TESTIS AND BREAST. LOWER LEVELS IN THYMUS,  
 CC LUNG, SALIVARY GLAND, AND PANCREAS.  
 CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE  
 CC RESIDUES.  
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: L07868; AAB59446.1; -.  
 DR HSSP: P11362; IFGR.  
 DR Genew: HGNC:3432; ERBB4.  
 DR MIM: 600543; -.  
 DR InterPro: IPR000494; EGFR\_L.domain.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR002174; Furin-like.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR InterPro: IPR004019; YLP\_motif.  
 DR Pfam: PF00069; pkinase.1.  
 DR Pfam: PF00757; Furin-like.1.  
 DR Pfam: PF01030; Recep\_L.domain.2.  
 DR Pfam: PF02757; YLP.2.  
 DR ProDom: PD000001; Euk\_pkinase.1.  
 DR SMART: SM00261; FU.4.  
 DR SMART: SM00219; TYRKC.1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP.1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR.1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM.1.  
 KW Transmembrane; Glycoprotein; Multigene family; Receptor; signal;

KW	Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
FT	Alternative splicing.
FT	FT SIGNAL 1 25
FT	FT CHAIN 26 1308
FT	FT DOMAIN 26 651
FT	FT TRANSMEM 652 675
FT	FT DOMAIN 676 1308
FT	FT DOMAIN 186 334
FT	FT DOMAIN 496 633
FT	FT DOMAIN 718 985
FT	FT NP_BIND 724 732
FT	FT BINDING 751 751
FT	FT ACT_SITE 843 843
FT	FT DISULFID 189 197
FT	FT DISULFID 193 205
FT	FT DISULFID 213 221
FT	FT DISULFID 217 229
FT	FT DISULFID 230 238
FT	FT DISULFID 234 246
FT	FT DISULFID 249 258
FT	FT DISULFID 262 289
FT	FT DISULFID 293 304
FT	FT DISULFID 308 323
FT	FT DISULFID 326 330
FT	FT DISULFID 303 312
FT	FT DISULFID 507 520
FT	FT DISULFID 523 532
FT	FT DISULFID 536 552
FT	FT DISULFID 555 569
FT	FT DISULFID 559 577
FT	FT DISULFID 580 589
FT	FT DISULFID 593 614
FT	FT DISULFID 617 625
FT	FT DISULFID 621 633
FT	FT MOD_RES 1162 1162
FT	FT MOD_RES 1186 1186
FT	FT MOD_RES 1258 1258
FT	FT MOD_RES 1284 1284
FT	FT CARBOHYD 138 138
FT	FT CARBOHYD 174 174
FT	FT CARBOHYD 181 181
FT	FT CARBOHYD 253 253
FT	FT CARBOHYD 358 358
FT	FT CARBOHYD 410 410
FT	FT CARBOHYD 473 473
FT	FT CARBOHYD 495 495
FT	FT CARBOHYD 548 548
FT	FT CARBOHYD 576 576
FT	FT CARBOHYD 620 620
FT	FT VARSPLIC 626 648
SO	SEQUENCE 1308 AA; 146807 MW; 5E4A8E095D88761 CnC64;
Query Match	36.7%; Score 1451; DB 1; Length 1308;
Best Local Similarity	43.5%; Pred. No. 1; le-102;
Matches 275; Conservativeness 99; Mismatches 238; Indels 20; Gaps 8;	
QY	9 WGLLALLPGAA-----STVCCTGDMKRLRLPASPERHLMRLHYGCGVCGNLELTY 64
DB	8 WWWSLVLAAGTYVPPSSQSYACETENKLSLSLEDOYRALRKRYINCEYVAGNLEITS 67
QY	65 LPTNASLFLDIOEVGYVLIANNOYVPLQRLIRVGTQLEFEDNYALAVLNDGDPIN 124
DB	68 IEHNRDLSPLRVREYGVYVLANGRVPLERLRIIRGKLYEDRYALALFLNRYKDG 127
QY	125 NTPPYTASPGGLRELDLRSITELKAGVLIQRNPOLCYQDTITMKILFKNNALTLI 164
DB	128 NF-----GLQELGKLNLTETLINGVYVDKFKFLADTTIMODIVRNPPSNLTIV 178
QY	185 DTNRBACHPCSPMCKGRCGSGESSEDCOSLTRVACGGC-ARCKGGLPTDCHEGCAAG 243
DB	179 STNSSGCGRCHSCTG-RCMGPTENHCOTILTRVCAEDCGKRYGTVSDCCRCACAG 237







DR ProDom: PD000001; Euk\_pkinase; 1.  
 DR SMART: SM00261; Fu; 3.  
 DR SMART: SM00219; TyrKc; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; FALSE\_NEG.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; FALSE\_NEG.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;  
 KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;  
 KW Alternative splicing.  
 FT SIGNAL 19  
 FT CHAIN 20  
 FT DOMAIN 20  
 FT TRANSSEM 644  
 FT DOMAIN 665  
 FT DOMAIN 709  
 FT NP\_BIND 715  
 FT BINDING 742  
 FT ACN\_SITE 834  
 FT DISULFID 186  
 FT DISULFID 190  
 FT DISULFID 210  
 FT DISULFID 227  
 FT DISULFID 231  
 FT DISULFID 246  
 FT DISULFID 259  
 FT DISULFID 290  
 FT DISULFID 305  
 FT DISULFID 323  
 FT DISULFID 327  
 FT DISULFID 500  
 FT DISULFID 504  
 FT DISULFID 520  
 FT DISULFID 532  
 FT DISULFID 533  
 FT DISULFID 552  
 FT DISULFID 556  
 FT DISULFID 576  
 FT DISULFID 589  
 FT DISULFID 613  
 FT DISULFID 617  
 FT CARBOHYD 126  
 FT CARBOHYD 250  
 FT CARBOHYD 353  
 FT CARBOHYD 408  
 FT CARBOHYD 414  
 FT CARBOHYD 437  
 FT CARBOHYD 469  
 FT CARBOHYD 522  
 FT CARBOHYD 566  
 FT CARBOHYD 616  
 FT VARSPLIC 141  
 FT  
 FT VARSPLIC 184  
 FT CONFLICT 560  
 FT CONFLICT 1064  
 FT SEQUENCE 1342 AA; 148097 MW; 7201E7F66CA37ABD CRC64;  
 Query Match 36.3%; Score 1435; DB 1; Length 1342;  
 Best Local Similarity 43.6%; Pred. No. 1.8e-10;  
 Matches 278; Conservative 92; Mismatches 244; Indels 24; Gaps 12;

QY 188 RSACHPSCPMCKSGKRCWGESEDCOSLTRTYVCAGGC-ARCKGPLYPTDCCHQCAAGCTG 246  
 DB 179 NRSRCPPEHVEVKG-KCMGPGSEDCQTLTKTTCARQCNHCGPBNQCCHPCAGGCGG 237  
 QY 247 PRHSDCLACLHFNHSGICELCPALVTYNTDFEESMPNDEGRYTFGASCVPYNYLST 306  
 DB 238 PDOTDCAFNRHNDGACVPRCPOLVYVKNLFTQLEPNHTKYQGVGVASCPIHNFV-V 296  
 QY 307 DVGSCITVCPRLHNOEYTAEDGTQRCCKSKPCARVCYGLGMHLEVRAYTSANIQERAG 366  
 DB 297 DQTSVCRAAPPCKMEVD-KNGLKMCPCGGLCPKACEGSGS-SRFQVDSNNIDGFVN 353  
 QY 367 CKRIGSLAFLEESPDGPPASNTAPLOPQLOVFETLEITGYLYISAMPDLPDLSYRQ 426  
 DB 354 CKRIIGNDELFTLGLNDGPMHRIKIPALDPKLVNFTVREITGYLIQSNPPHMHFVS 413  
 QY 427 NLQVIRGRILHNGAYS-LTLQGLISWGLRSRLREIGSGLALIHNTHLCEVHTVPMQDL 485  
 DB 414 NLTTIGGRSLYKRGSLIMKLNLYTSIGFSLKEISAGRIYISANRQLCYHNSLWTKV 473  
 QY 486 FNPHQALHTA-NRPEDECVGEGLACQOLCARGCMWGPYQCVNCSOFLMGQCEVEEC 544  
 DB 474 IKGPEERLDIKHNPBRDCVAEGKVCPLCSSGCGWGPQCLSCRMYSRGVCVYTHC 533  
 QY 545 RVLOGLPREYVARHCLPCHPECPONGSVTCFGEADQVCACAHYKDPFCVAPCPGV 604  
 DB 534 NFLNDEPRFEHAEACFSCHPECPMEGTATNGSSSDTCAQCAHFRDGFHCYSSCPHGV 593  
 QY 605 KPDLSYMPKFPDEGACQPCPINCETHSCVDLDDKGC 642  
 DB 594 LG-AKGPYKYPDVQNECRPCHEMCTOGCKRPELDQC 629

RESULT 10  
 ERB3\_RAT  
 ID 062799; STANDARD: PRT; 1339 AA.  
 AC 15-DEC-1998 (Rel. 37, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Receptor protein-tyrosine kinase erbB-3 precursor (EC 2.7.1.112)  
 DE (c-erbB3).  
 GN ERB3.  
 OS Rattus norvegicus (rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NCBI\_TaxID=10116;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Sprague-Dawley; TISSUE-Liver;  
 RX MEDLINE=96096535; PubMed=8522190;  
 RA Hellyer N.J., Kim H.-H., Greaves C.H., Sierke S.L., Roland J.G.;  
 RT "Cloning of the rat ErbB3 cDNA and characterization of the  
 RT recombinant protein.";  
 RL Gene 165:279-284(1995).  
 [2]  
 RN REVISIONS TO 85; 513 AND 565.  
 RP Hellyer N.J., Roland J.G.;  
 RL Submitted (DEC-2001) to the EMBL/Genbank/DBJ databases.  
 [3]  
 RP SEQUENCE OF 922-1097 FROM N.A.  
 RC STRAIN-Sprague-Dawley; TISSUE-Sciatic nerve;  
 RX MEDLINE=97184212; PubMed=9030624;  
 RA Carroll S.L., Miller M.L., Frohnert P.W., Kim S.S., Corbett J.A.;  
 RT "Expression of neurotrophins and their putative receptors, ErbB2 and  
 RT ErbB3, is induced during Wallerian degeneration.";  
 RL J. Neurosci. 17:1642-1659(1997).  
 CC -1- FUNCTION: BINDS AND IS ACTIVATED BY NEUREGULINS AND NTAF.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS  
 CC (POTENTIAL).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- DOMAIN: THE CYTOPLASMIC PART OF THE RECEPTOR MAY INTERACT WITH THE  
 CC SH2 OR SH3 DOMAINS OF MANY SIGNAL-TRANSDUCING PROTEINS.  
 CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE RESIDUES  
 CC AND PROMOTES ITS ASSOCIATION WITH THE P85 SUBUNIT OF  
 CC PHOSPHATIDYLINOSITOL 3-KINASE.  
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U09339; AAC38498.2; -;  
 DR EMBL: U05530; AAC53050.1; -;  
 DR HSSP: P11362; 1FGK;  
 DR InterPro: IPR000494; EGFR\_L\_domain.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR002174; Furin-like.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR Pfam: PF00069; pkinase.1.  
 DR Pfam: PF00757; Furin-like.1.  
 DR Pfam: PF01030; Recep\_L\_domain.2.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR ProDom: PD000001; Euk\_pkinase.1.  
 DR SMART: SM00261; Fuv.5  
 DR SMART: SM00219; TyrKc.1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; FALSE\_NEG.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR.1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM.1.  
 DR Transmembrane: Glycoprotein; Multiase family; Receptor; Signal;  
 KW Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.  
 FT SIGNAL 1 19  
 FT CHAIN 20 1339  
 FT DOMAIN 20 643  
 FT TRANSMEM 644 662  
 FT DOMAIN 663 1339  
 FT DOMAIN 183 259  
 FT NP\_BIND 707 964  
 FT BINDING 713 721  
 FT ACT\_SITE 740 740  
 FT ACT\_SITE 832 832  
 FT DISULFID 186 194  
 FT DISULFID 190 202  
 FT DISULFID 210 218  
 FT DISULFID 214 226  
 FT DISULFID 227 235  
 FT DISULFID 231 243  
 FT DISULFID 246 255  
 FT DISULFID 259 286  
 FT DISULFID 290 301  
 FT DISULFID 305 320  
 FT DISULFID 323 327  
 FT DISULFID 327 327  
 FT DISULFID 504 517  
 FT DISULFID 520 529  
 FT DISULFID 533 549  
 FT DISULFID 556 573  
 FT DISULFID 576 585  
 FT DISULFID 589 610  
 FT DISULFID 613 621  
 FT DISULFID 617 629  
 FT CARBOHYD 126 126  
 FT CARBOHYD 250 250  
 FT CARBOHYD 353 353  
 FT CARBOHYD 408 408  
 FT CARBOHYD 414 414  
 FT CARBOHYD 437 437  
 FT CARBOHYD 469 469  
 FT CARBOHYD 522 522  
 FT CARBOHYD 566 566

FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 1028 1028 L -> P (IN REF. 3).  
 SQ SEQUENCE 1339 AA: 147545 MW: 0A5F2402BBDFLE CXC64;  
 Query Match 34.8%; Score 1374.5; DB 1; Length 1339;  
 Best Local Similarity 42.6%; Pred. No. 7.4e-97;  
 Matches 272; Conservative 87; Mismatches 250; Indels 29; Gaps 13;  
 QY 3 LALACRGGLLALLPPAA---SYOVCTGDMKRLRPASETHIDMLRHLYXGCGVQGN 59  
 DB 7 LQVLC---FLSLARSEMGNSOAVCPGTLNGSLVGDADNOYQTLKYECEVVMGN 62  
 QY 60 LEIIVLPYLTNASLSELDIOIEVQGVLLIAHNOVQVPLQRLRIYRGTLFEDNYALVDN 119  
 DB 63 LEIIVLGHNDLSFLQWIREVTGVIVANNEFSLPRLNRYVRGTVGYGKALIFM-- 120  
 QY 120 GDLINNTPYVTGASPGGLRELQRLSLEILKGVLIQRPOLCYODTILMKDIFHNKQL 179  
 DB 121 ---LNYNT---NSSHALROLKFTOLTEILSGGYIEKNKRLCHMDTIDRDIYVR-- 170  
 QY 180 ALTLIDFNRSRACHPCSPMGKSGRSGWSESDOSLFRVYCAAGG-ARCKGPLPTDCHE 238  
 DB 171 GAELIVAKNNGANCPRCHVEVKG-RCWGEGPDQOILKTLICAPQCNGRCGPPMNOCCHD 229  
 QY 239 QCAAGCTGPKHSDCLACIAHNSGICELHCPALVYNTDTFESMPNPEGRTFGASCVTA 298  
 DB 230 ECAGGCGSPQDDQCFACRRRNDSGACVPRDEPLVYKLFQLEPNHTKYQYGGCVAS 289  
 QY 299 CPYVYLTSDVSGCLVCPRLNDEYTAEDGTORCEKSGPCARVCYGLGMEHLEVRVYS 358  
 DB 290 CPNHFV-VDQTFVCRACPDPKMEVD-NHGLKMEPCGGLCPKAGEGSGS--SRQYVDS 345  
 QY 359 ANIOEFAKCKIKFSLAFLEPSFDGDPASNTAPLPOLQVFETLEITGVLYIASPDS 418  
 DB 346 SNIDGFNCKRIKGLNDFILITGLNVDPMKRIPALDEPKLVFRVRRITGVLNQSMPH 405  
 QY 419 LPDLVSFQNLQVIRGRILHNGAYS-LTLQGLIGISWGLSLRELGSGLALIHNTLCFV 477  
 DB 406 MHNFSVSNLTIGGRSLYNGFSLIMKMLNVTSLGFSRLKETSAGRVYISANOQCYH 465  
 QY 478 HTVWMDLFFNPQALLHTA-NRPEDECVGEGLACHOLCARGHGWGPGVQVNSCFRL 536  
 DB 466 HSLMTWRLLEGPSEERLDIYDRPLGLECLAEKVCYDPLSSGCGWGPQOJCSRNYSR 525  
 QY 537 GQECVEECRVLYGLPREYVNAHCLPCHPCOPQNSVTFEGPEADQVCANHKDPPFC 596  
 DB 526 EGVVTHCNFLQGRPREFVIEAQCFSCHPEDCLPMEGISTCNGSSSDACARCAHRODPHC 585  
 QY 597 VAPCPGVPKPDLSYMPIWKFPEDEGACQPCPINCTHSC 634  
 DB 586 VNSCPHGILG-AKGPYKYKPDADQNECRPCHECTQGC 621  
 RESULT 11  
 XMRK\_XIPMA STANDARD; PRT; 1167 AA.  
 ID XMRK\_XIPMA  
 AC P1338;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Melanoma receptor protein-tyrosine kinase precursor (EC 2.7.1.112).  
 GN XMRK OR TU.  
 OS Xiphophorus maculatus (Southern platyfish).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Altieriomorpha;  
 OC Cyprinodontiformes; Poeciliidae; Xiphophorus.  
 OX NCBI\_Taxid=8083;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90015140; PubMed=2797166;  
 RA Witbrodt J., Adam D., Malitschek B., Maueier W., Raulf F.,  
 RA Telling A., Robertson S.M., Scharf M.

RT "Novel putative receptor tyrosine kinase encoded by the melanoma-inducing tu locus in Xiphophorus.";  
 RL Nature 341:415-421(1989).  
 RN [2]  
 RP REVISION TO 515.  
 RA Schartl M.;  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: PROBABLE RECEPTOR WITH PROTEIN-TYROSINE KINASE ACTIVITY.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein tyrosine phosphate.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- DISEASE: INVOLVED IN PIGMENT CELLS MALIGNANT MELANOMAS.  
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/announce/or\\_send\\_an\\_email\\_to\\_license@sib-sib.ch](http://www.isb-sib.ch/announce/or_send_an_email_to_license@sib-sib.ch)).  
 CC -----  
 DR EMBL: X16691, CA34770.2; -;  
 DR PIR: S06142, S06142.  
 DR HSSP: P11362, IFGK.  
 DR Interpro: IPR000494; EGFR\_Ldomain.  
 DR Interpro: IPR000719; Euk\_pkinase.  
 DR Interpro: IPR002174; Furin-like.  
 DR Interpro: IPR002290; Ser\_thr\_pkinase.  
 DR Interpro: IPR001245; Tyr\_pkinase.  
 DR Pfam: PF00757; Furin-like; 1.  
 DR Pfam: PF00757; Furin-like; 1.  
 DR Pfam: PF00757; Furin-like; 1.  
 DR PRINTS: PR00109; Recpt\_Ldomain; 2.  
 DR ProDom: PD000001; Euk\_pkinase; 1.  
 DR SMART: SM00261; FU; 5.  
 DR SMART: SM00220; S\_TKC; 1.  
 DR SMART: SM00219; TYRK; 1.  
 DR SMART: SM00219; TYRK; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 KW Transmembrane; Glycoprotein; Receptor; Signal; Transferase;  
 KW Tyrosine-protein kinase; ATP-binding; Phosphorylation; Proto-oncogene.  
 FT SIGNAL 1 25  
 FT \* CHAIN 26 1167 MELANOMA RECEPTOR PROTEIN-TYROSINE KINASE.  
 FT DOMAIN 26 642  
 FT TRANSMEM 643 665 EXTRACELLULAR (POTENTIAL).  
 FT DOMAIN 666 1167 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 710 977 PROTEIN KINASE.  
 FT NP\_BIND 716 724 ATP (BY SIMILARITY).  
 FT BINDING 743 743 ATP (BY SIMILARITY).  
 FT ACT\_SITE 835 835  
 FT DISULFID 195 204 BY SIMILARITY.  
 FT DISULFID 199 212 BY SIMILARITY.  
 FT DISULFID 220 228 BY SIMILARITY.  
 FT DISULFID 224 236 BY SIMILARITY.  
 FT DISULFID 237 245 BY SIMILARITY.  
 FT DISULFID 241 253 BY SIMILARITY.  
 FT DISULFID 256 265 BY SIMILARITY.  
 FT DISULFID 269 296 BY SIMILARITY.  
 FT DISULFID 300 311 BY SIMILARITY.  
 FT DISULFID 315 330 BY SIMILARITY.  
 FT DISULFID 333 337 BY SIMILARITY.  
 FT DISULFID 504 513 BY SIMILARITY.  
 FT DISULFID 508 521 BY SIMILARITY.  
 FT DISULFID 524 533 BY SIMILARITY.  
 FT DISULFID 537 553 BY SIMILARITY.  
 FT DISULFID 556 569 BY SIMILARITY.  
 FT DISULFID 560 577 BY SIMILARITY.  
 FT DISULFID 593 615 BY SIMILARITY.  
 FT DISULFID 618 626 BY SIMILARITY.  
 FT DISULFID 622 634 BY SIMILARITY.

FT CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 201 201 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 365 365 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 417 417 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 501 501 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 576 576 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 621 621 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 1167 AA; 129934 MW; 4793E0749DC1D35A CRC64;  
 Query Match 33.6%; Score 1330; DB 1; Length 1167;  
 Best Local Similarity 41.7%; Pred. 1.6e-93;  
 Matches 269; Conservative 93; Mismatches 259; Indels 24; Gaps 11;  
 QY 4 AALCFMGLLALLPPGAASV---OYCTGTDMLRLPASPEHMLDMLRLHYOGCOVQGN 59  
 DB 8 AALLQ--LLLVLSIRCSCTDPDRKVCQGTSTNQMTM---LDNHVLMKKMYSGCNVYLEN 62  
 QY 60 LEITLPTNASTFTLQDIOEVGYVLIANOVROYPLORLRYRGTOLEFEDNYALAVDN 119  
 DB 63 LEITTYQENQDLSFQDSIQEVGYVLIANNEVSTPLRLRIGONLYEGNFTLVMSN 122  
 QY 120 GDPNNTPVTGASPGSLREQLRSITELTKGVLIQRPOLCYODTILMKDIFPKNNQ 179  
 DB 123 YK--NPSSP--DYVQGLKQDLSNLTLSGVAVSHNPILCANVETIMWDIVKTSNP 179  
 QY 180 ALTLIDTNKSRACHPCSPCKSRGWSSESDQSLRTVPCAGG--ARCKGLPTDCHE 238  
 DB 180 TKNLIPHAERQCKQCDHCGVNGSCWAPGRCQKFTKLCAEQGNRRRGKRPIDCCNE 239  
 QY 239 QCAAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDTEFSMPNPEBRTFFGASCTA 298  
 DB 240 HCAGGCTGPRATDCLACRFNDGTCKDPCPKIYDYSRQVYNPNRYTFFGAACYKE 299  
 QY 299 CPYNTLSTPVGSCVTLCPPLHNOEYTAEDGTQCEKSCSKCAVAGYGLGMEHLREVRAYS 358  
 DB 300 CENNVYVE--GACVASCASGMLFVD--ENGRSKCKRDGVCPRVCGIGISLNTIAVNS 357  
 QY 359 ANIOFAGCKKIFGSLAFIPESFGDPPASNTAPLOPOLQVETLEETLYGLYISAMPDS 418  
 DB 358 TNIRSFNSCKTKINGDILINRNSFEGBDPHYKICTMPDEHLMNLTUYKEITGYLVIMWPN 417  
 QY 419 LPDLVFPQNLQYRGRILHNGAYS--LTLOGLSLMLGLRSLBELSGLALIHNNHLCFV 477  
 DB 418 MSLSVFQNLLEIRRTFTSRGSEFVVQVRHQLMGLRSLKESVAGVILNLTQLRYA 477  
 QY 478 HTVPMDOLFRRNHQALLHTANRPEDCEVGEGLACHQOLARGCHWGPPTOCVNSQPLRG 537  
 DB 478 NTIINRRLEFRSDQSIETDART-----ENQTCNNECSDEGCMGPRPTMVCSLHVDRG 530  
 QY 538 QECVECRVLOGLPREYVNAHRLCPHPECOPONGSVTCFGEADQCAVAAHYKDPPECV 597  
 DB 531 GRCVASCNLLQEPREDAQDRCVQHOECLVQTSILTCYGRGANCSSAHPQGPQCI 590  
 QY 598 ARCPGSGVADLSYMPIMKFPPEDEGACQCPRTICTSHCYVDLDDKGC 642  
 DB 591 PRCPHGILGDDTL--IMKYADKMGQCPCHQNCCTGCGSGPGLSGC 634  
 RESULT 12  
 EGF\_R\_DROME STANDARD; PRT: 1426 AA.  
 AC P04412; O61601; Q9W2G0; P81868;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Epidermal growth factor receptor precursor (EC 2.7.1.112) (Egfr)  
 DE (Gurken receptor) (torpedo protein) (Drosophila relative of ERBB).  
 GN EGFR OR TOP OR C-ERBB OR DER OR CG10079.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;

OC Insecta; pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS TYPES I AND II).  
 RX MEDLINE=94350209; PubMed=8070664;  
 RA Clifford R., Schupbach T.;  
 RT "Molecular analysis of the Drosophila EGF receptor homolog reveals  
 RT that several genetically defined classes of alleles cluster in  
 RT subdomains of the receptor protein.";  
 RL Genetics 137:531-550(1994).  
 RN [2]  
 RP REVISIONS.  
 RA Clifford R., Schupbach T.;  
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
 RX MEDLINE=85124611; PubMed=2982499;  
 RA Livneh E., Glazer L., Segal D., Schlessinger J., Shilo B.-Z.;  
 RT "The Drosophila EGF receptor gene homolog: conservation of both  
 RT hormone binding and kinase domains.";  
 RL Cell 40:599-607(1985).  
 RN [4]  
 RP SEQUENCE FROM N.A., CHARACTERIZATION, AND ALTERNATIVE SPLICING.  
 RC STRAIN=Oregon R.; TISSUE=Embryo;  
 RX MEDLINE=87002474; PubMed=3093080;  
 RA Schejter E.D., Segal D., Glazer L., Shilo B.-Z.;  
 RT "Alternative 5' exons and tissue-specific expression of the  
 RT Drosophila EGF receptor homolog transcripts.";  
 RL Cell 46:1091-1101(1986).  
 RN [5]  
 RP SEQUENCE FROM N.A. (ISOFORM TYPE I), TISSUE SPECIFICITY, AND MUTATION  
 RP ANALYSIS.  
 RX MEDLINE=99102120; PubMed=9882502;  
 RA Lesochin A.M., Yu S.-Y., Katz J., Baker N.E.;  
 RT "Several levels of EGF receptor signaling during photoreceptor  
 RT specification in wild-type, Ellipse, and null mutant Drosophila.";  
 RL Dev. Biol. 205:129-144(1999).  
 RN [6]  
 RP SEQUENCE FROM N.A. (ISOFORM TYPE I).  
 RC STRAIN=Berkely;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Ceoliker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Mortan J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazel R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Planck C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
 RA Burtis K.C., Busan D.A., Butler H., Cadiou E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson R., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ideguchi J.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kemison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muray D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleby J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Massarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 RN [7]  
 RP SEQUENCE OF 959-1078 FROM N.A.  
 RC STRAIN=Daekwanryeong;  
 RX MEDLINE=85137938; PubMed=2983232;  
 RA Wadsworth S.C., Vincent W.S. III, Bilodeau-Wentworth D.;  
 RT "A Drosophila genomic sequence with homology to human epidermal  
 RT growth factor receptor.";  
 RL Nature 314:178-180(1985).  
 RN [8]  
 RP SEQUENCE OF 1133-1137; 1155-1167 AND 1209-1216 FROM N.A., AND MUTATION  
 RP ANALYSIS.  
 RX MEDLINE=92038942; PubMed=1936959;  
 RA Raz E., Schejter E.D., Shilo B.-Z.;  
 RT "Interallelic complementation among DER/Elb alleles: implications for  
 RT the mechanism of signal transduction by receptor-tyrosine kinases.";  
 RL Genetics 129:191-201(1991).  
 RN [9]  
 RP REVIEW.  
 RX MEDLINE=97248481; PubMed=9094709;  
 RA Perittom N., Perkins L.A.;  
 RT "There must be 50 ways to rule the signal: the case of the Drosophila  
 RT EGF receptor.";  
 RL Cell 89:13-16(1997).  
 CC -I- FUNCTION: BINDS TO FOUR LIGANDS: SPITZ, GURKEN, VEIN AND ARGOS,  
 CC WHICH IS AN ANTAGONIST. TRANSDUCES THE SIGNAL THROUGH THE RAS-RAF-  
 CC MAPK PATHWAY. INVOLVED IN A TRIAD OF DEVELOPMENTAL DECISIONS.  
 CC CRITICAL FOR THE PROLIFERATION OF IMAGINAL TISSUES, AND FOR THE  
 CC DETERMINATION OF BOTH THE ANTERO-POSTERIOR AND DORSO-VENTRAL  
 CC POLARITIES OF THE OOCYTE. IN THE EMBRYO, PLAYS A ROLE IN THE  
 CC ESTABLISHMENT OF VENTRAL CELL FATES, MAINTENANCE OF AMNIOSERA  
 CC AND VENTRAL NEUROECTODERMAL CELLS, GERM BAND RETRACTION, CELL FATE  
 CC SPECIFICATION IN THE CENTRAL NERVOUS SYSTEM AND PRODUCTION OF  
 CC CUTICLE.  
 CC -I- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -I- SUBCELLULAR LOCATION: ISOFORMS TYPE I AND TYPE II ARE TYPE I  
 CC MEMBRANE PROTEINS. ISOFORM TYPE III IS A TYPE III MEMBRANE  
 CC PROTEIN.  
 CC -I- ALTERNATIVE PRODUCTS: 3 ISOFORMS; TYPE I (SHOWN HERE), TYPE II AND  
 CC TYPE III; ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -I- TISSUE SPECIFICITY: UBQUITOUSLY EXPRESSED IN EMBRYOS. IN LARVAE,  
 CC UNIFORM EXPRESSION IS SEEN IN WING DISKS, GENITAL DISK, ANLAGEN OF  
 CC TESTIS AND OVARY, AND BRAIN CORTEX. IN EYE-ANTENNA DISK, HIGHEST  
 CC EXPRESSION IS ANTERIOR TO MORPHOGENETIC FURROW, LEVELS REMAIN HIGH  
 CC IN PHOTORECEPTOR PRECURSOR CELLS. THIS PATTERN IS REVERSED IN  
 CC POSTERIOR EYE DISK. IN ADULTS EXPRESSION IS HIGH IN BRAIN CORTEX  
 CC AND THORACIC AND ABDOMINAL GANGLIA.  
 CC -I- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation - its  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF052754; AAC08536.1; -;  
 CC EMBL: AF052753; AAC08536.1; JOINED.  
 CC EMBL: AF052754; AAC08535.1; -;  
 CC EMBL: AF052752; AAC08535.1; JOINED.  
 CC EMBL: K03054; AAA51462.1; -;  
 CC EMBL: K03417; AAA51460.1; -;  
 CC EMBL: K03416; AAA50965.1; -;  
 CC EMBL: K03418; AAA51461.1; -;  
 CC EMBL: AF109077; AAD26134.1; -;  
 CC

```

DR EMBL; AF109078; AAD26132.1; -.
DR EMBL; AF109082; AAD26132.1; JOINED.
DR EMBL; AF109078; AAD26133.1; -.
DR EMBL; AF109084; AAD26133.1; JOINED.
DR EMBL; AF109079; AAD26130.1; -.
DR EMBL; AF109081; AAD26130.1; JOINED.
DR EMBL; AF109079; AAD26131.1; -.
DR EMBL; AF109083; AAD26131.1; JOINED.
DR EMBL; AF109080; AAD26135.1; -.
DR EMBL; AF003454; AAF46732.1; -.
DR EMBL; X02293; CAA26157.1; -.
DR EMBL; X78920; CAA55523.1; -.
DR EMBL; X78918; CAA55522.1; -.
DR EMBL; X78919; CAA55522.1; -.
DR PIR; A00640; GGFPE.
DR HSP; P11362; IFGK.
DR FLYBase; FBgn0003731; Egfr.
DR InterPro; IPR000494; EGFR_L_domain.
DR InterPro; IPR000719; Euk_Pkinase.
DR InterPro; IPR002174; Furin-like.
DR InterPro; IPR001245; Tyr_Pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF01030; Recep_L_domain; 2.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_Pkinase; 1.
DR SMART; SM00261; FU; 7.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
DR Transmembrane; Glycoprotein; Receptor; Phosphorylation; Transferase;
KW Tyrosine-protein kinase; ATP-binding; Signal; Alternative splicing;
KW Developmental protein.
FT SIGNAL 1 30 POTENTIAL.
FT CHAIN 31 1426 EPIDERMAL GROWTH FACTOR RECEPTOR.
FT DOMAIN 31 868 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 869 889 POTENTIAL.
FT DOMAIN 890 1426 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 938 1198 PROTEIN KINASE.
FT NP_BIND 944 952 ATP (BY SIMILARITY).
FT BINDING 971 971 ATP (BY SIMILARITY).
FT ACT_SITE 1063 1063 BY SIMILARITY.
FT MOD_RES 902 902 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).

Query Match 25.68; Score 1013.5; DB 1; length 1426;
Best Local Similarity 33.08; Pred. No. 2.6e-69;
Matches 231; Conservative 93; Mismatches 280; Indels 97; Gaps 20;

QY 24 QVCTGTDKRLPASPETHDMLRHLYGCOVQGNLELTVLP-NMSTSLDIOEVQ 82
DB 100 KICGTSRSLVSPSKENHYNLNDRTNCTYVGNLKLTVLPENLDLSTLDNIREVTG 159
QY 83 YVLAHNOVQVPLQRLRIVRGTLF-----EDNYALAVLDNGDPLNNTTFTVTSAPGL 137
DB 160 YILSHVDVKVPPKLIQIINGRTLFSLVEEKKYALFV-----TYSKM 203
QY 138 RELDLRSLTELKGVLIQRNPQLCYODTILMKDIFKKNOLALTLDITNSRACHPCSP 197
DB 204 YTLLEIPDLRLDVLNGOVGFHNNYNLCMHMTIOWSEIVSNGTDAYNYDTABERDECPKHE 263
QY 198 MCKSRRCGSESEDCQSLTRTVAGGCA--RCKGPLPTDCHEQCAAGCTGPKHSDCLAC 255
DB 264 SCTHG-CMGEPRKNCQKFSKLTCSPOCAGGRCYGPKRPECHLFCAGGCTPTQKDCIAC 322
QY 256 LAFNHSGICELHCPALVYTYNTDFESMPNPGRYTFGASCVTACPYNYLSTDVGSCTLVC 315
DB 323 KNFDEAVSKRECPMRKYNPTTVLETFNPEGYAVGATVCYKCP-GHLLDNGACVASC 381
QY 316 PLHNDVTAEBGTORCEKSPCARVCYGLGMEHLREVRANTSANIQEFACCKIIFGLA 375
DB 382 PQDKMDKGGF-----CVPKNGPCPKTCPGVTLH-----AGNIDSFRCNTVIDGNIR 428

```

```

QY 376 FLPSFDS--DPASNTA-----PLQPROLOVEFTEETFGTLYISAMPDLSPLDSFVN 427
DB 429 ILDTFSGFQDYVANYTYMGPRYIPLDPERREVEFSTVKETITGTLNIEGHPPFRNLSTFRN 488
QY 428 LQVIRGRILHNGAY-SLTLOGISWLGRLSRLRELSGLALIHNTHLCPVHTYPMWDLF 486
DB 489 LETINGRLMESMPFAALAIYVKSLSLTLSEMRNLKQISSSSVYIQHNRDLCYSNTIMPAIQ 548
QY 487 RNPQALLHTANRPDECVSGSLACHQLCARGHCWPGPPTQVCNCSQFLRQCEVEECRV 546
DB 549 KEPEQKVVWVENLRADLLEKNGTICSDQCNEDGCGAGTDCOLCKNFENGTCTIADCGY 608
QY 547 LOGLPREVNARHCLPCHPEQOPQNGSVTCGPREDDCVACAHKDPFPCVAPRCPSSGVP 606
DB 609 ISMAK--FDNRKCKICHPECR-----TCNAGADHCQECVHVNDGQHCVCSECKN-- 657
QY 607 DLSTYPIWKPFDEEGACPCPINCTHSCVDLDDKCPCEORASPLTSQNEIDLGPASPL-- 664
DB 658 -----KYND-RGVCECHATC-----DGC-----TGPKDITIGIACITC 690
QY 665 -----DSTFYRLLEDDMDGLVDAEYLVPO-QGFPCP 697
DB 691 NLATINNDATVVKRCILKDKCPDGY-FWEYVHPQEGSLKP 730

RESULT 13
ID LT23_CAEEL STANDARD; PRT; 1323 AA.
AC P24348;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE let-23 receptor protein-tyrosine kinase precursor (EC 2.7.1.112).
GN let-23 OR KIN-7 OR ZK1067.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
R1 SEQUENCE FROM N.A.
R2 MEDLINE=9108019; Pubmed=1979659;
R3 Aroian R.V., Koga M., Mendel J.E., Ohshima Y., Sternberg P.W.;
R4 "The let-23 gene necessary for Caenorhabditis elegans vulval
R5 induction encodes a tyrosine kinase of the EGF receptor subfamily.";
R6 Nature 348:693-699(1990).
RL [2]
R1 SEQUENCE FROM N.A.
R2 STRAIN=Bristol N2;
R3 MEDLINE=9617760; Pubmed=8604137;
R4 Sakai T., Koga M., Ohshima Y.;
R5 "Genomic structure and 5' regulatory regions of the let-23 gene in
R6 the nematode C. elegans.";
R7 J. Mol. Biol. 256:548-555(1996).
RL [3]
R1 SEQUENCE FROM N.A.
R2 STRAIN=Bristol N2;
R3 Thomas K.;
R4 submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
RL [4]
R1 MUTANTS.
R2 MEDLINE=94147981; Pubmed=8313880;
R3 Aroian R.V., Les G.M., Sternberg P.W.;
R4 "Mutations in the Caenorhabditis elegans let-23 EGFR-like gene define
R5 elements important for cell-type specificity and function.";
R6 EMBO J. 13:360-366(1994).
RL [5]
R1 SUBCELLULAR LOCATION.
R2 STRAIN=Bristol N2;
R3 MEDLINE=99287744; Pubmed=10359617;
R4 Whitfield C.W., Benard C., Barnes T., Hekimi S., Kim S.K.;
R5 "Basolateral localization of the Caenorhabditis elegans epidermal
R6 growth factor receptor in epithelial cells by the p23 protein
R7 LIN-10.";

```



RL Mol. Biol. Cell 10:2087-2100(1999).  
CC - FUNCTION: Tyrosine kinase receptor required for the induction of  
CC vulval differentiation. Possible receptor for the inductive signal  
CC required for vulval development. Activated by lin-3 and acts by  
CC way of let-60 Ras. The lin-3/let-23 pair is a simplified version  
CC of the mammalian neuroligin-ERBB network.  
CC - CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
CC tyrosine phosphate.  
CC - SUBCELLULAR LOCATION: Type I membrane protein. Basolateral and  
CC apical membrane of cell junctions in epithelial vulval precursor  
CC cells.  
CC - TISSUE SPECIFICITY: Vulval precursor cells.  
CC - DEVELOPMENTAL STAGE: Expressed during L2 and L3 larval stages.  
CC - SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: X57767; CAA00919.1; ALT-SEQ.  
DR EMBL: D63426; BAA09729.1; ALT-INIT.  
DR EMBL: Z70038; CAA93882.1; -  
DR PIR: S13422; S13422.  
DR HSSP: P11362; 1FGK.  
DR WormPep: zk1067.1; CE03840.  
DR InterPro: IPR000494; EGRF\_L\_domain.  
DR InterPro: IPR000719; Euk\_pkinase.  
DR InterPro: IPR002174; Furin-like.  
DR InterPro: IPR001245; Tyr\_pkinase.  
DR Pfam: PF00069; pkinase; 1.  
DR Pfam: PF00757; Furin-like; 1.  
DR Pfam: PF01030; Recept\_L\_domain; 2.  
DR ProDom: PD000001; Euk\_pkinase; 1.  
DR SMART: SM00261; FG; 6.  
DR SMART: SM00219; TYRKC; 1.  
DR ProSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR ProSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR ProSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR Transfaser: Tyrosine-protein kinase; Developmental protein;  
KW ATP-binding; Phosphorylation; Transmembrane; Glycoprotein; Receptor;  
KW signal.  
FT SIGNAL. 1 22  
FT CHAIN 23 1333 LEFT-23 RECEPTOR PROTEIN-TYROSINE KINASE.  
FT DOMAIN 23 818 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 819 841 POTENTIAL.  
FT DOMAIN 842 1323 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 845 1152 PROTEIN KINASE.  
FT NP\_BIND 891 899 ATP (BY SIMILARITY).  
FT BINDING 919 919 ATP (BY SIMILARITY).  
FT ACT\_SITE 1010 1010 BY SIMILARITY.  
FT DISULFID 220 226 BY SIMILARITY.  
FT DISULFID 224 236 BY SIMILARITY.  
FT DISULFID 244 251 BY SIMILARITY.  
FT DISULFID 248 262 BY SIMILARITY.  
FT DISULFID 263 271 BY SIMILARITY.  
FT DISULFID 267 279 BY SIMILARITY.  
FT DISULFID 282 291 BY SIMILARITY.  
FT DISULFID 295 322 BY SIMILARITY.  
FT DISULFID 326 337 BY SIMILARITY.  
FT DISULFID 341 356 BY SIMILARITY.  
FT DISULFID 359 364 BY SIMILARITY.  
FT DISULFID 320 329 BY SIMILARITY.  
FT DISULFID 524 537 BY SIMILARITY.  
FT DISULFID 540 549 BY SIMILARITY.  
FT DISULFID 553 567 BY SIMILARITY.  
FT DISULFID 570 577 BY SIMILARITY.  
FT DISULFID 574 585 BY SIMILARITY.  
FT DISULFID 588 604 BY SIMILARITY.  
FT DISULFID 608 620 BY SIMILARITY.

FT DISULFID 623 632 BY SIMILARITY.  
FT DISULFID 627 644 BY SIMILARITY.  
FT DISULFID 647 660 POTENTIAL.  
FT DISULFID 670 693 POTENTIAL.  
FT DISULFID 696 703 POTENTIAL.  
FT DISULFID 700 715 POTENTIAL.  
FT DISULFID 717 731 POTENTIAL.  
FT DISULFID 735 750 POTENTIAL.  
FT DISULFID 753 763 POTENTIAL.  
FT DISULFID 757 771 POTENTIAL.  
FT DISULFID 774 787 POTENTIAL.  
FT DISULFID 791 805 POTENTIAL.  
FT CARBOHYD 91 91 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 169 169 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 376 376 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 561 561 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 635 635 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 746 746 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 776 776 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT VARIANT 368 368 C -> Y (IN SY10).  
FT VARIANT 469 469 C -> R (IN MN216).  
FT VARIANT 700 700 C -> W (IN MN23).  
FT VARIANT 733 733 C -> Y (IN SY11).  
FT VARIANT 1065 1065 T -> I (IN SY16).  
FT VARIANT 1074 1074 G -> E (IN SY7).  
SQ SEQUENCE 1323 AA; 150510 MW; 6B0307E853EEFA99 CRC64;

Query Match 18.2%; Score 718.5; DB 1; Length 1323;  
Best Local Similarity 28.0%; Pred. No. 7.2e-47;  
Matches 191; Conservative 107; Mismatches 256; Indels 129; Gaps 28;

QY 25 VCTGTGTMKRLPASPETHLDMRLHLYOGCOVGNELRYLPNT----- 68  
DB 39 LCSGTINGISRYTGNI-LEDLETMTRGCRRYVGNLEFIWENETKKMRSTNVDR 97  
QY 69 -----ASLSFLDIOEVOGYVLIHANOVRYPLORLIRVGTQLEFDNYALAVLDNGDP 122  
DB 98 NEDSPLKSTINFPNLEIRSLIIVYANIKISFRLRYVYGEVPHDN-ALYIHRNDR- 155  
QY 123 LANTTVYTGASPGGLKLELRSLTELKGVLIQRNPOLCY-ODTILMKDIFKNNQAL 181  
DB 156 -----VHEVYRELVRIRNGSVTTIQDNPKNKYIDKIDMKWELVDPD--VO 199  
QY 182 TLIDFNRSRACH-----PCSPMGKGRGCGESSEDECSLFRVYAGGACARC---KGPL 231  
DB 200 KVETTSHQHCYONGKSMKACHESC--NDKCGSGMDNDQVRVYVCPKSCSGCFYNSIS 258  
QY 232 PTDCHEQCAAGCTGPRHSDCLAFHNHSGICELCPALVYNTDTFESMPNPEGRYTF 291  
DB 259 SYECDSACAGCGCTGHPKNCIACSKYELDGCIEFCSKRIRNHRGRLVFNPDGRYON 318  
QY 292 GASCYACAPNYL-STDVGSCTLYC-PLINQEVTAEDGQRCRC-SKRCARCYGLGME 348  
DB 319 GNHCVKECPPELLIENDV--CYRHCSGDGHYDATKP--VRECEKCRSSSCPCKICTYDVG-- 372  
QY 349 HIREVAVYTSANIOERFAGCKKIRGSLAFIPESDPGDPASTAVLQEOLOVPEITBEIG 408  
DB 373 HL-----INETLKNLEGCEQIDGHL-IIEHA-----TEQLKLVETVAVISB 414  
QY 409 YLYISAMPDLPDLVSFONLOVIRGILHNHGAVALTL-QLGISWLGSLRSLRELGSGLAL 467  
DB 415 YITIV--QGNFYDLKRLKNQIIEGRKLNHVRALAIYQCDDELSLMSIKIKGAVL 472  
QY 468 IHNHTLCTFYHYVPWOLF-----RNPHOALLTANRPEDECGEGGLACHQLCARHGCKP 523  
DB 473 IMKNHRLCYVSKIDWSIITSKGDMPKSLATAENRDSKICETEORVCPCKNCKRCCKW 532  
QY 524 GPTQCVNCSQFLGSCGVECEKRVLOGLPREYVARHCLCHPCPCQONQNSVYCFGEADQ 583  
DB 533 EPEDCLECTKWSVGTCEVKCDI-KGFLRNQTSMK-CERCSPCE-----TGNGLGEID 584  
QY 584 CVACAH-----YKDPPECVARGPSGVRPDLSTYMPIMKFPPEGACQPCPINCHSCVD 636

```

DB 565 CLTCHKRLKLYNSDFGNRMECHDC-----PVSHFPRQKRVCEKCHPTCY----- 628
OY 637 LDDKCCPAEDQASPLTSONEDLG 659
DB 629 --DNCC-----TGPDSNLG 640

RESULT 14
ID ILPR_BRAVA STANDARD; PRT; 1363 AA.
AC 002466;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Insulin-like peptide receptor precursor (EC 2.7.1.112) (ILP receptor).
OS Branchiostoma lanceolatum (Common lancelet) (Amphioxus).
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
OC Branchiostoma.
OC NCBI_Taxid=7740;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96408719; PubMed=8813726;
RA Pashmforoush M., Chan S.J., Steiner D.F.;
RT "Structure and expression of the insulin-like peptide receptor from
RT amphioxus."
RL Mol. Endocrinol. 10:857-866(1996).
CC -1- FUNCTION: THIS RECEPTOR BINDS TO THE INSULIN RELATED PEPTIDE AND
CC HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- SUBUNIT: PROBABLE Tetramer of 2 alpha and 2 beta chains linked by
CC disulfide bonds. The alpha chains contribute to the formation of
CC the ligand-binding domain, while the beta chain carry the kinase
CC domain (by similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (by similarity).
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. INSULIN
CC RECEPTOR SUBFAMILY.
CC -----
CC * THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S83394; AAB50848.1; -.
DR HSSP; P06213; IIRK.
DR InterPro; IPR000494; EGFR_L_domain.
DR InterPro; IPR000719; Euk_Pkinase.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR003962; FNIII_repeat.
DR InterPro; IPR002174; Furin-like.
DR InterPro; IPR002011; RTkinaseI.
DR InterPro; IPR001245; Tyr_Pkinase.
DR Pfam; PF00041; fn3.3.
DR Pfam; PF00069; Pkinase.1.
DR Pfam; PF00757; Furin-like.1.
DR Pfam; PF01030; Recep_L_domain.2.
DR PRINTS; PR00014; FNTYPEIII.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Euk_Pkinase.1.
DR SMART; SM00060; FN3.3.
DR SMART; SM00261; FU.1.
DR SMART; SM00219; TYRK.1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP.1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR.1.
DR PROSITE; PS00239; RECEPTOR_TYR_KIN_III.1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM.1.
DR Transferrase; Tyrosine-protein kinase; Receptor; Transmembrane;
KW Glycoprotein; ATP-binding; Phosphorylation; Signal.
FT SIGNAL 1 29 POTENTIAL.

```

```

FT CHAIN 30 716 INSULIN-LIKE PEPTIDE RECEPTOR, ALPHA-
FT PROPEP 717 720 CHAIN (POTENTIAL).
FT CHAIN 721 1363 POTENTIAL.
FT DOMAIN 721 928 INSULIN-LIKE PEPTIDE RECEPTOR, BETA-
FT TRANSMEM 929 949 CHAIN (POTENTIAL).
FT DOMAIN 950 1363 EXTRACELLULAR (POTENTIAL).
FT NP_BIND 994 1283 POTENTIAL.
FT BINDING 1000 1008 CYTOPLASMIC (POTENTIAL).
FT ACT_SITE 1028 1028 PROTEIN KINASE.
FT MOD_RES 1148 1174 ATP (BY SIMILARITY).
FT CARBOHYD 51 51 BY SIMILARITY.
FT CARBOHYD 97 97 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CARBOHYD 137 137 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 278 278 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 483 483 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 599 599 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 617 617 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 665 665 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 666 666 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 711 711 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 732 732 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 736 736 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 743 743 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 816 816 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 885 885 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 898 898 N-LINKED (GLCNAc...) (POTENTIAL).
SQ SEQUENCE 1363 AA; 154104 MW; 238120B4EAB1ED65 CRC64;

```

Query Match 9.7%; Score 384; DB 1; Length 1363;  
 Best Local Similarity 26.3%; Pred. No. 2,3e-21;  
 Matches 137; Conservative 56; Mismatches 180; Indels 148; Gaps 25;

```

OY 9 WGL-----LALLPFGASTGYCTGTDKRLRASPETHLDRLHYGGCQVQGNLELT 63
DB 12 WAALLVYIGLGLVNSNEEYICDSMDIRN-----VSLRLD-ENCYIIEGYLDI- 61
OY 64 YLPTNASLFLDIOEVGYVLIHNOYQV-----LQRLR-----IVRGTO 106
DB 62 -----LLIDFAEEDYSGIAPNVEIYDFLYLRVGLPMISELPMIAVIRGTN 112
OY 107 LFEDNYVALVNDGPNLNTTPVTGASPGGLRELOLRSTELKGCYLQNRPOLCYODT 166
DB 113 LF-FNYALVFEMLD-----WQKIGLYSLQNTTRGSVRIEKNPNLCYIDT 156
OY 167 ILMKDIF---HKNNQALATLIDTNRSRAC-HPCSPMK-----GSRGWGESSEDCQSILT 216
DB 157 IDMSFIAESGYSNN-----FIYDNRREECVNCPCRCRIKRVYLDLQMAE--EHGQKRV 210
OY 217 RTVCAGGCAKCGKPLPTDCHEQCAAGCTGPKHSCCLACLHNHSGIEELHCPALVYNT 276
DB 211 PESCIQNR---DSISGCHEHCIGCGDGPTEPCVACKFVHNGECLICPPPTQYK 266
OY 277 D-----TFESMPREGRY--TFGASCVTAQPYNYVISTDVGSCITVQPLANQETATADGQR 330
DB 267 DRCITTEECPTTNSVVKLHHRKCIPECPSTY-PTDINNPEL----- 308
OY 331 CEKSKPCARVCYGLGEMLEVRRAVTSANIOEFAGCKKIFSLAFLESPEDGASMTA 390
DB 309 CTECGQCPKSCGGLVSL-----AAQRFNGCIIIBELKISRIGD----- 352
OY 391 PLQPEQLQVFTLEETGYL-----YISAMPD-SLPDLSVFQNTQVIRGRIILHNGAYSILT 444
DB 353 -----NIIDLEENLGLIEVGHVAIVRSYALVTDLFSLKRIRIGIQENG-YAFY 404
OY 445 LQGLISWLGSLREL-----GSGLALIHNNHLC 475
DB 405 V-----LDNRNLEKLPDMQDRTITIDEGKLFFHFNPRLC 438

RESULT 15

```

ID	IR_R_MOUSE	STANDARD:	PRT:	1300 AA.
AC	IR_R_MOUSE			
AC	OSWt14:			
DT	15-JUN-2002 (Rel. 41, Created)			
DT	15-JUN-2002 (Rel. 41, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Insulin receptor-related protein precursor (EC 2.7.1.112) (IRR)			
DE	(IRR-related receptor).			
GN	INSR OR IRR			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_Taxid=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99272238; PubMed=10342810;			
RA	Hirayama I., Tanemoto H., Yokota H., Kubo S.-K., Wang J., Kuwano H.,			
RA	Nagamachi Y., Takeuchi T., Izumi T.;			
RT	"Insulin receptor-related receptor is expressed in pancreatic b-cells			
RT	and stimulates tyrosine phosphorylation of insulin receptor			
RT	substrate-1 and-2."			
RL	Diabetes 48:1237-1244(1999).			
CC	-I- FUNCTION: This receptor probably binds an insulin related protein			
CC	and has a tyrosine-protein kinase activity. It phosphorylates the			
CC	insulin receptor substrates IRS-1 and IRS-2.			
CC	-I- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein			
CC	tyrosine phosphate.			
CC	-I- SUBUNIT: PROBABLE TETRAMER OF 2 ALPHA AND 2 BETA CHAINS LINKED BY			
CC	DISULFIDE BONDS. THE ALPHA CHAINS CONTRIBUTE TO THE FORMATION OF			
CC	THE LIGAND-BINDING DOMAIN, WHILE THE BETA CHAIN CARRY THE KINASE			
CC	DOMAIN (By similarity).			
CC	-I- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).			
CC	-I- TISSUE SPECIFICITY: Highly expressed in the islets as well as in			
CC	pancreatic beta-cells.			
CC	-I- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. INSULIN			
CC	RECEPTOR SUPERFAMILY.			
CC	-I- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE III-LIKE DOMAINS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; AB007135; BAA7835.1; -			
DR	HSSP; P06213; IIRK.			
DR	MCD; MGI:1346037; Insir.			
DR	InterPro: IPR000494; EGFR_L_domain.			
DR	InterPro: IPR000719; Euk_pkinase.			
DR	InterPro: IPR003961; FN_III.			
DR	InterPro: IPR002174; Furin-like.			
DR	InterPro: IPR002011; RTkinasel.			
DR	InterPro: IPR001245; Tyr_pkinase.			
DR	Pfam; PF00041; fn3; 2.			
DR	Pfam; PF00069; pkinase; 1.			
DR	Pfam; PF00757; Furin-like; 1.			
DR	Pfam; PF01030; Recep_L_domain; 2.			
DR	PRINTS; PR00109; TYRKINASE.			
DR	ProDom; PD000001; Euk_pkinase; 1.			
DR	SMART; SMO0060; FN3; 3.			
DR	SMART; SMO0261; FU; 1.			
DR	SMART; SMO0219; TyfKC; 1.			
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.			
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.			
DR	PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.			
DR	PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.			
KW	Transferrase; Tyrosine-protein kinase; Receptor; Transmembrane;			
KW	Glycoprotein; ATP-binding; Phosphorylation; Repeat; Signal.			
FT	SIGNAL	1	26	POTENTIAL.
FT	CHAIN	27	1300	INSULIN RECEPTOR-RELATED PROTEIN, ALPHA-
FT	CHAIN	27	746	INSULIN RECEPTOR-RELATED PROTEIN, ALPHA-

Mon Jan 13 17:52:22 2003

us-09-854-356-7.rsp

Page 21

Oy 431 IGRHLNGAYSL-----TLOGLGISNL--GLRSRLRELSGALLHNHTLCEYH 478  
||| : : ||| : ||| : ||| :  
Db 402 IRGSNWDGNVTLVLDNQNLQQLG-SWVTAGL----TIIVGRTVEAFNPRLCLEH 452

Search completed: January 13, 2003, 14:47:11  
Job time : 15.6039 secs

**THIS PAGE BLANK (uapfo)**